TANGENT LINES, INFLECTIONS, AND VERTICES OF CLOSED CURVES

MOHAMMAD GHOMI

Abstract. We show that every smooth closed curve Γ immersed in Euclidean space $\mathbf{R}^3$ satisfies the sharp inequality $2(\mathcal{P}+\mathcal{I})+\mathcal{V}\geq 6$ which relates the numbers $\mathcal{P}$ of pairs of parallel tangent lines, $\mathcal{I}$ of inflections (or points of vanishing curvature), and $\mathcal{V}$ of vertices (or points of vanishing torsion) of Γ. We also show that $2(\mathcal{P}^++\mathcal{I})+\mathcal{V}\geq 4$, where $\mathcal{P}^+$ is the number of pairs of concordant parallel tangent lines. The proofs, which employ curve shortening flow with surgery, are based on corresponding inequalities for the numbers of double points, singularities, and inflections of closed curves in the real projective plane $\mathbf{RP}^2$ and the sphere $\mathbf{S}^2$ which intersect every closed geodesic. These findings extend some classical results in curve theory including works of Möbius, Fenchel, and Segre, which is also known as Arnold's "tennis ball theorem".

Contents

Date: Last Typeset January 19, 2012.

1991 *Mathematics Subject Classification.* Primary: 53A04, 53C44; Secondary: 57R45, 58E10.

Key words and phrases. Inflection, vertex, torsion, flattening point, curve shortening, mean curvature flow, four vertex theorem, tennis ball theorem, singularity, skew loop, tangent line.

Research of the author was supported in part by NSF Grant DMS-0336455.

1. Introduction

To every point of a regular closed space curve $\gamma\colon \mathbf{S}^1 \simeq \mathbf{R}/2\pi \to \mathbf{R}^3$ there corresponds a matrix $m := (\gamma', \gamma'', \gamma''')$ composed of the first three derivatives of γ. Naturally one is interested in degeneracies of m, which come in two geometric flavors: *inflections* or points of vanishing curvature, and *vertices* or points of vanishing torsion. Here we show how inflections and vertices are related to the global arrangement of the tangent lines of γ, which leads us to extend some well known theorems of classical curve theory. To state our first result, let us say that a pair of parallel tangent lines of γ are *concordant* if the corresponding tangent vectors point in the same direction, and are *discordant* otherwise. A vertex is called *genuine* if it has a connected open neighborhood where the torsion, or equivalently $\det(m)$, vanishes just once and changes sign.

Theorem 1.1. *Let $\gamma\colon \mathbf{S}^1 \to \mathbf{R}^3$ be a regular $\mathcal{C}^3$ closed curve, and $\mathcal{P}$ be the number of pairs of distinct points in $\mathbf{S}^1$ where tangent lines of γ are parallel. Further let $\mathcal{I}$ and $\mathcal{V}$ be the numbers of inflections and vertices of γ respectively. Then*

$$2(\mathcal{P} + \mathcal{I}) + \mathcal{V} \geq 6. \tag{1}$$

Furthermore, if $\mathcal{P}^+ (\leq \mathcal{P})$ denotes the number of concordant *pairs of parallel tangent lines, then*

$$2(\mathcal{P}^+ + \mathcal{I}) + \mathcal{V} \geq 4. \tag{2}$$

Finally, if γ has only finitely many vertices, then $\mathcal{V}$ may denote the number of genuine *vertices.*

Thus if a closed space curve has nonvanishing curvature and torsion (which is a generic condition with respect to the $\mathcal{C}^1$-topology), then it must have at least three pairs of parallel tangent lines including at least two concordant pairs. This observation is meaningful because there exist closed space curves, called *skew loops*, without any pairs of parallel tangent lines. Skew loops, which were first studied by Segre [39], may be constructed in each knot class [49], and have been the subject of several recent works [15, 18, 16, 21, 20, 41] due in part to their interesting connections with quadric surfaces. We should also mention that (2) had been observed by Segre [39] in the special case where $\mathcal{P}^+ = \mathcal{I} = 0$, while (1) appears to be entirely new.

As is often the case in space curve theory, Theorem 1.1 follows from a reformulation of it in terms of spherical curves, which is of interest on its own. To derive this deeper result first note that, assuming γ has unit speed, the *tantrix* and *curvature* of γ are given by $T := \gamma'$, and $\kappa := \|T'\|$ respectively. Thus inflections of γ correspond precisely to *singularities* of $T\colon \mathbf{S}^1 \to \mathbf{S}^2$, or points where T' vanishes. Otherwise, the *geodesic curvature* of T in $\mathbf{S}^2$, given by

$$k := \frac{\langle T'', T \times T' \rangle}{\kappa^3} = \frac{\tau}{\kappa},$$

is well-defined, where τ is the torsion of γ. Thus vertices of γ correspond precisely to *(geodesic) inflections* of T, as is well known [13]. Further, the tangent lines of γ are parallel at t, $s \in \mathbf{S}^1$ provided that $T(t) = \pm T(s)$, and are concordant if $T(t) = T(s)$.

It is also well known that a spherical curve forms the tantrix of a space curve, if, and only if, it contains the origin of $\mathbf{R}^3$ in the relative interior of its convex hull [13, 17]. Via these observations, inequalities (1) and (2) follow respectively from inequalities (3) and (4) below. In continuing the analogy with Theorem 1.1, we say that an inflection is *genuine* if it has a connected open neighborhood where curvature vanishes just once and changes sign.

Theorem 1.2 (The Main Result)**.** *Let $\gamma\colon \mathbf{S}^1 \to \mathbf{S}^2$ be a $\mathcal{C}^2$ closed curve, $\mathcal{S}$ be the number of singular points of γ, $\mathcal{I}$ be the number of its inflections, and $\mathcal{D}$ be the number of pairs of points $t \neq s \in \mathbf{S}^1$ where $\gamma(t) = \pm\gamma(s)$. Suppose that $\Gamma := \gamma(\mathbf{S}^1)$ contains the origin in its convex hull. Then*

$$2(\mathcal{D} + \mathcal{S}) + \mathcal{I} \geq 6. \tag{3}$$

Further, if $\mathcal{D}^+(\leq \mathcal{D})$ denotes the number of pairs of points $t \neq s \in \mathbf{S}^1$ where $\gamma(t) = \gamma(s)$, then

$$2(\mathcal{D}^+ + \mathcal{S}) + \mathcal{I} \geq 4. \tag{4}$$

Furthermore, if Γ is symmetric*, i.e., $\Gamma = -\Gamma$, then*

$$2(\mathcal{D}^+ + \mathcal{S}) + \mathcal{I} \geq 6. \tag{5}$$

Finally, if γ has only finitely many inflections, then $\mathcal{I}$ may denote the number of genuine *inflections.*

Inequality (3) does not seem to have appeared before, not even in the special case where $\mathcal{D} = \mathcal{S} = 0$. Inequalities (4) and (5), on the other hand, extend some well known results. Indeed, when $\mathcal{D}^+ = \mathcal{S} = 0$, (4) is a theorem of Segre [40, 47], a special case of which was rediscovered by Arnold and famously dubbed "the tennis ball theorem" [5, 6, 3]. Further, when $\mathcal{S} = 0$, (4) follows from works of Fenchel [12] according to Weiner [47, Thm. 1]. Furthermore, when $\mathcal{D}^+ = \mathcal{S} = 0$, (5) is equivalent to a classical theorem of Möbius [31, 3, 42], who showed that simple closed noncontractible curves in the real projective plane $\mathbf{RP}^2 := \mathbf{S}^2/\{\pm 1\}$ have at least 3 inflections. More generally, Theorem 1.2 has the following quick implication for curves in $\mathbf{RP}^2$:

Theorem 1.3. *Let $\gamma\colon \mathbf{S}^1 \to \mathbf{RP}^2$ be a $\mathcal{C}^2$ closed curve, $\mathcal{S}$ be the number of singular points of γ, $\mathcal{I}$ be the number of its inflections, and $\mathcal{D}$ be the number of pairs of points $t \neq s \in \mathbf{S}^1$ where $\gamma(t) = \gamma(s)$. If γ is noncontractible, then*

$$2(\mathcal{D} + \mathcal{S}) + \mathcal{I} \geq 3. \tag{6}$$

On the other hand, if γ is contractible, and its image intersects every closed geodesic, then

$$2(\mathcal{D} + \mathcal{S}) + \mathcal{I} \geq 6. \tag{7}$$

Proof. If γ is noncontractible, then there is a closed symmetric curve $\overline{\gamma}\colon \mathbf{S}^1 \to \mathbf{S}^2$ which double covers $\gamma(\mathbf{S}^1)$ when composed with the standard projection $\pi\colon \mathbf{S}^2 \to \mathbf{RP}^2$. So (5) immediately implies (6). On the other hand, if γ is contractible, then there is a closed curve $\overline{\gamma}\colon \mathbf{S}^1 \to \mathbf{S}^2$ such that $\pi \circ \overline{\gamma} = \gamma$. Thus $\overline{\gamma}$ has the same number

of inflections and singularities as γ has, since π is a local isometry. Further, double points of γ correspond to the self intersections of $\overline{\gamma}(\mathbf{S}^1)$, and to the intersections of $\overline{\gamma}(\mathbf{S}^1)$ with $-\overline{\gamma}(\mathbf{S}^1)$. So (3) implies (7) once we note that if $\gamma(\mathbf{S}^1)$ intersects every closed geodesic in $\mathbf{RP}^2$, then $\overline{\gamma}(\mathbf{S}^1)$ does the same in $\mathbf{S}^2$, and thus contains the origin in its convex hull. $\square$

The rest of this work will be devoted to proving Theorem 1.2, which will be achieved through a series of reductions. First, in Section 3, we use some basic convexity theory and a lemma for spherical curves in the spirit of Fenchel [12] to prove Theorem 1.2 in the special case where $\mathcal{D}^+ = \mathcal{S} = 0$ and Γ lies in a hemisphere. This generalizes a result of Jackson [26], also studied by Osserman [33], which is an extension of the classical 4-vertex-theorems of Mukhopadhyaya [32] and Kneser [30] for planar curves (Note 3.4). Next, in Section 4, we employ the curve shortening flow after Angenent [3] and Grayson [22] to show that in the case where $\mathcal{D}^+ = \mathcal{S} = 0$, we may continuously transform Γ until it just lies in a hemisphere. This transformation does not increase any of the quantities $\mathcal{D}$, $\mathcal{D}^+$, $\mathcal{I}$ and $\mathcal{S}$, and thus leads to a proof of Theorem 1.2 in the case $\mathcal{D}^+ = \mathcal{S} = 0$, via the hemispherical case already considered in Section 3. In particular, we obtain new proofs of Möbius's and Segre's theorems at this stage. Then it suffices to show that one may remove the double points and singularities of Γ without increasing the sums $2(\mathcal{D} + \mathcal{S}) + \mathcal{I}$ or $2(\mathcal{D}^+ + \mathcal{S}) + \mathcal{I}$, and while keeping the origin in the convex hull of Γ. The desingularization and surgery procedures which we need here will be developed in Sections 5 and 6. Then in Section 7 we will be able to synthesize these results to complete the proof. Finally we construct a number of examples in Section 8 which will demonstrate the optimality of all the above inequalities.

The study of inflections and vertices, which dates back to Klein's estimate for algebraic curves [28] and Möbius's 3-inflections-theorem, has spawned a vast literature of interesting though often isolated results. One bright focal point, however, has been Segre's theorem, which implies the works of Mukhopadhyaya and Kneser (Note 2.5), as well as that of Möbius (Lemma 4.6). It also leads to more recent results such as Sedykh's four-vertex-theorem for space curves (Note 1.5), and has connections to contact geometry, as noted by Arnold [5, 6]. For other recent developments and references see [19, 14, 8, 34, 43, 35] and note that vertices are also known as *flattening points* [4, 45]. Some open problems will be discussed in Notes 1.5, 8.1, and 8.3.

Note 1.4. Inequalities (3) and (4) for spherical curves are somewhat more general than their counterparts, inequalities (1) and (2), for space curves. Indeed, as we mentioned earlier, the tantrix of a space curve will always contain the origin in the relative *interior* of its convex hull [17], whereas in Theorem 1.2 the origin may lie on the boundary of the convex hull. These critical cases require special care throughout this work.

Note 1.5. The term "vertex" is used in the literature to refer not only to zero torsion points of space curves, but also to local extrema of the geodesic curvature of a curve in a Riemannian surface. This terminology is motivated by the fact that

for spherical curves critical points of geodesic curvature correspond to points of vanishing torsion. Indeed if p is a critical point of the geodesic curvature of $\Gamma \subset \mathbf{S}^2$, then Γ has contact of order 3 with its osculating circle C in $\mathbf{S}^2$. But C coincides with the osculating circle of Γ as a curve in $\mathbf{R}^3$. Thus Γ has contact of order 3 with its osculating plane at p, and so has zero torsion at that point. The double use of the term vertex may be further justified by Sedykh's theorem [38] which states that if a simple closed space curve lies on the boundary of its convex hull, then it must have (at least) four vertices (zero torsion points). This is analogous to Mukhopadhyaya's classical four-vertex theorem for convex planar curves. See the paper of Thorbergsson and Umehara [42] for a simple proof of Sedykh's result based on Segre's theorem. We should also mention a related open problem due to Rosenberg [36]: *must every closed space curve bounding a surface of positive curvature have four vertices?* This question is related to a problem of Yau [50, Problem 26] on characterizing boundaries of positively curved surfaces, which in turn is motivated by rigidity questions for closed surfaces in $\mathbf{R}^3$. See [1] for more references in this area.

2. Preliminaries

We begin by developing our basic terminology, and recording some fundamental observations which will be useful throughout the paper.

2.1. **Basic terminology.** Here $\mathbf{R}^n$ denotes the n-dimensional Euclidean space with origin o, standard inner product $\langle \cdot, \cdot \rangle$, and norm $\| \cdot \| := \langle \cdot, \cdot \rangle^{1/2}$. The (closed) unit ball and sphere in $\mathbf{R}^n$ will be denoted by B^n and $\mathbf{S}^{n-1} := \partial B^n$ respectively. We often set $B := B^2$, or more generally let B denote a metric ball (or disk) in a Riemannian surface. Let I denote a connected one-dimensional manifold, which for convenience we identify either with an interval in $\mathbf{R}$, or $\mathbf{S}^1 \simeq \mathbf{R}/2\pi$. A *curve* is a locally injective mapping $\gamma\colon I \to M$, where M is some manifold. If γ is injective, then we say that it is *simple*. A pair of curves γ_1, γ_2 are considered *topologically equivalent* if there is a homeomorphism $\theta\colon I \to I$ such that $\gamma_2 = \gamma_1 \circ \theta$. A point $t_0 \in I$ is called a *singular point* of γ provided that the differential map $\gamma' := d\gamma$ is either not well-defined or is singular at t_0. Otherwise t_0 will be called a *regular point* of γ. Accordingly we say that $p \in M$ is a *regular value* of γ provided that every point of $\gamma^{-1}(p)$ is regular; otherwise, p is a *singular value*.

It will often be convenient to represent a class of topologically equivalent curves $[\gamma]$ by its image. More precisely, we say that $\Gamma \subset M$ is a curve provided that there is a fixed equivalence class of curves $\gamma\colon I \to M$ associated to Γ such that $\gamma(I) = \Gamma$. The elements of $[\gamma]$ will then be called *the (admissible) parametrizations* of Γ. We say that Γ is $\mathcal{C}^k$ if it admits a $\mathcal{C}^k$ parametrization; furthermore, if this parametrization is regular, then we say that Γ is $\mathcal{C}^k$*-regular* (or $\mathcal{C}^k$-immersed). We say Γ is *simple* if it has a simple parametrization. If a point p of a $\mathcal{C}^k$ curve Γ is a singular value of all $\mathcal{C}^k$ parametrizations of it, then p is called a ($\mathcal{C}^k$) *singular point* of Γ; otherwise, p is a *regular point* of Γ. If $\gamma^{-1}(p)$ is a singleton, we say that p is a *simple point* of Γ; otherwise p will be called a *multiple point*, and if $\gamma^{-1}(p)$ consists

of precisely two points, then p will be a *double point*. Finally, we say that Γ is an *arc* when $I \subset \mathbf{R}$ is a compact interval, and Γ is *closed* when $I = \mathbf{S}^1$.

2.2. **The maximum principle.** If Γ is a $\mathcal{C}^2$-regular curve in a Riemannian surface M, and N is a unit normal vector field along Γ, then the *geodesic curvature* of Γ at a point $p \in \Gamma$ with respect to N is defined as

$$k(p) := \big\langle \nabla_{T(p)} T, N(p) \big\rangle,$$

where T is a (continuous) unit tangent vector field along Γ, ∇ is the covariant derivative, and $\langle \cdot, \cdot \rangle$ denotes the metric on M. We say that p is an *inflection* of Γ if $k(p) = 0$, and Γ is a *geodesic* if k vanishes identically.

By *near* a point, we mean in a sufficiently small open neighborhood of that point. Let $\Gamma \subset M$ be a simple $\mathcal{C}^1$-regular curve, p be an interior point of Γ, and $B \subset M$ be a metric ball centered at p. If Γ is an arc or a closed curve, and B is sufficiently small, then Γ meets ∂B only at its end points. Thus $B - \Gamma$ will consist of precisely two components, by Jordan's curve theorem. The closure of each of these components will be called a *side* of Γ near p. The following basic fact, which is a version of the maximum principle, will be invoked a number times in the following pages.

Lemma 2.1. *Let M be a Riemannian surface with nonnegative curvature, and $\Gamma \subset M$ be a $\mathcal{C}^1$-regular simple arc which is $\mathcal{C}^2$-regular in its interior. Further let k be the geodesic curvature of Γ in its interior with respect to a unit normal vector field N along Γ, and C be a geodesic arc which is tangent to Γ at an end point p. Suppose that $N(p)$ points to the side of C which contains (resp. does not contain) Γ near p. Then either Γ coincides with C near p, or else $k > 0$ (resp. < 0) at some point in every neighborhood of p.*

Proof. Let $\pi\colon \Gamma \to C$ be the nearest point projection. After replacing Γ with a smaller subarc pq we may assume that π is well-defined (by the tubular neighborhood theorem) and one-to-one (by the inverse function theorem), since Γ is tangent to C at p. So we may assume that Γ is a "graph" over C. Further, we may assume that Γ lies in a metric ball $B \subset M$ centered at p, meets ∂B only at q, and C is also a simple arc in B which meets ∂B only at its end points. There is then a component of $C - \{p\}$ which together with Γ forms a simple $\mathcal{C}^1$-regular arc, say $\overline{\Gamma}$, which lies in B and intersects ∂B only at its end point; see Figure 1. By assumption C lies on one side of $\overline{\Gamma}$ in B. Let N be the unit normal vector field along Γ which points into the side of $\overline{\Gamma}$ containing C. Then $N(p)$ points to the side of C which does not contain Γ.

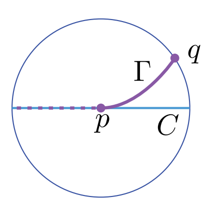


Figure 1.

Suppose that Γ does not coincide with C near p. Then we may assume that $q \notin C$. Let $C_\epsilon \subset B$ be the curve at the distance ϵ away from C which lies on the same side of C as Γ; see Figure 2. Then C_ϵ intersects pq for sufficiently small $\epsilon > 0$. Let q' be the first point of Γ, as we traverse it starting from p, which lies on C_ϵ, and p' (which may coincide with p) be the last point in pq' which lies on C. Now joining the arc $p'q'$ to the geodesic arc between q' and $\pi(q')$, and the geodesic arc between $\pi(q')$ and p', we obtain a closed curve, which is simple since Γ is a graph over C. Let $\Omega \subset B$ be the region bounded by this curve. Then the exterior angles of $\partial\Omega$ at p', $\pi(q')$, and q', are respectively π, $\pi/2$, and $\pi/2 + \delta$ for some $\delta \geq 0$.

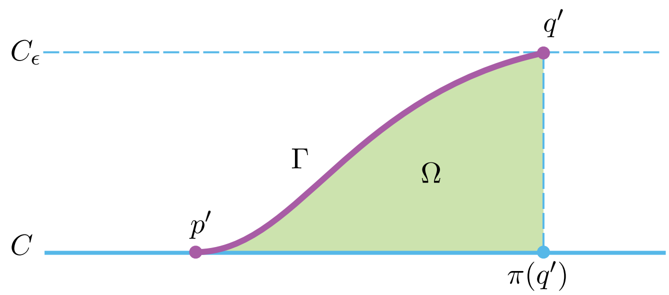


FIGURE 2.

Our choice of N ensures that N points inside Ω along the interior of $p'q'$. Thus applying the Gauss-Bonnet theorem to Ω yields:

$$\int_{p'}^{q'} k + \int_\Omega K + 2\pi + \delta = 2\pi, \tag{8}$$

where K is the Gauss curvature of M. So, since by assumption $K \geq 0$, it follows that $\int_{p'}^{q'} k \leq 0$. Thus either $k \equiv 0$ on the interior of $p'q'$ or else $k < 0$ somewhere between p' and q'. Since $q' \notin C$ and $p'q'$ is tangent to C at p', k cannot vanish identically on $p'q'$ by the uniqueness of geodesics. So we conclude that $k < 0$ somewhere on $p'q'$ which is the desired result, since q' may be chosen arbitrarily close to p.

Finally note that if we replace N with the opposite unit normal vector field along $\overline{\Gamma}$ (so that $N(p)$ points to the side of C containing Γ), then the geodesic curvature of the interior of $p'q'$ with respect to the inward normal of Ω will be $-k$. Replacing k with $-k$ in (8) yields that $k > 0$ somewhere on $p'q'$, by repeating the argument in the last paragraph. $\square$

2.3. **Convex spherical sets.** A set $X \subset \mathbf{S}^2$ is called *convex* if every pair of points of X may be joined by a geodesic arc of length at most π which lies in X. It is easy to see that $X \subset \mathbf{S}^2$ is convex if, and only if, the cone in $\mathbf{R}^3$ generated by X is convex. Thus many basic facts in the theory of convex Euclidean sets have direct analogues for convex spherical sets. In particular, the closure of any convex spherical set is convex. We need the following local characterization for closed convex spherical curves, i.e., curves which bound a convex set.

Lemma 2.2. *Let $\Gamma \subset \mathbf{S}^2$ be a simple closed curve bounding a region Ω. Suppose that for every point $p \in \Gamma$, at least one of the following conditions hold:*

(i) *There exists a geodesic arc passing through p which is disjoint from the interior of Ω.*
(ii) *There exists an open neighborhood U of p in Γ such that U is $\mathcal{C}^2$-regular, and the geodesic curvature of U is nonnegative with respect to the normal vector field pointing into Ω.*
(iii) *There is no geodesic arc in Ω which passes through p.*

Then Γ is convex.

The above lemma has well known direct analogues for planar curves, e.g., see [10]. Indeed, this lemma follows from its planar version in the case where Ω lies in an open hemisphere, via the Beltrami mapping discussed in the next section; however, we cannot a priori assume that Ω is confined to a hemisphere, and therefore an independent proof is needed.

Proof of Lemma 2.2. It suffices to show that the interior of Ω, which we denote by $\mathrm{int}(\Omega)$, is convex, since as we remarked above the closure of a convex spherical set is convex. Let x_0, $x_1 \in \mathrm{int}(\Omega)$. Since $\mathrm{int}(\Omega)$ is connected and open, it is path connected. So there is a continuous map $x\colon [0,1] \to \mathrm{int}(\Omega)$ such that $x(0) = x_0$, and $x(1) = x_1$. For every $t \in [0,1]$, let G_t be a geodesic arc of length at most π which connects $x(0)$ and $x(t)$. Now let A be the set of all $t \in (0,1]$ such that $G_t \subset \mathrm{int}(\Omega)$. Clearly A is nonempty and open. We claim that A is also closed, which will complete the proof. To see this let $t_i \in A$ be a sequence of points converging to $t \in (0,1]$. We have to show that then $G_t \subset \mathrm{int}(\Omega)$.

First we show that $G_t \subset \Omega$. To see this let $B \subset \mathrm{int}(\Omega)$ be a ball centered at $x(t)$, which does not contain $x(0)$; see Figure 3. Then for i sufficiently large, $G_i := G_{t_i}$ intersects ∂B at a point y_i. Since ∂B is compact, after replacing y_i with a subsequence, we may assume that y_i converge to $y \in \partial B$. Note that y cannot be antipodal to $x(0)$, because the lengths of G_i approach the distance between y and $x(0)$ plus the radius of B. Thus if the distance between $x(0)$ and y were π, then the lengths of G_i would eventually exceed π, which is not permitted. So we may assume

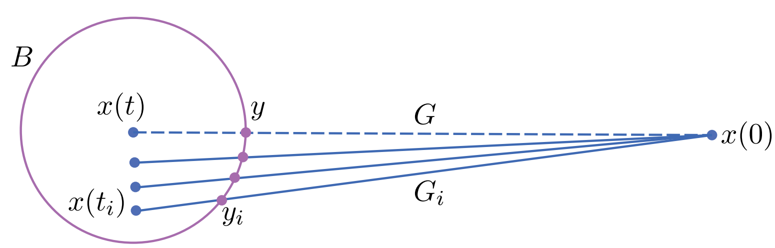


FIGURE 3.

that, for i sufficiently large, there are unique great circles C_i passing through $x(0)$ and y_i. Consequently C_i, and the arcs $G_i \subset C_i$, depend continuously on y_i. Let C be the great circle passing through $x(0)$ and y. Then C_i converge to C in the sense

of Hausdorff distance, i.e., C_i eventually lie inside any given open neighborhood of C. Consequently C passes through $x(t)$, since C_i contain $x(t_i)$ which converge to $x(t)$. Now let G be the arc of C which passes through y and is bounded by $x(0)$ and $x(t)$. Then G_i converge to G. Thus, since Ω is compact, $G \subset \Omega$. Furthermore, since G_i have length at most π, then so does G. Thus $G_t = G \subset \Omega$, as desired.

Now to show that $G_t \subset \mathrm{int}(\Omega)$, it suffices to check that G_t is disjoint from $\partial\Omega = \Gamma$. Suppose that G_t touches Γ at some point p. We claim then that $G_t \subset \Gamma$ which is the desired contradiction, since the end points of G_t lie in $\mathrm{int}(\Omega)$. To establish this claim let $X := G_t \cap \Gamma$. Then X is a nonempty closed subset of G_t. It suffices then to show that X is also open in G_t. To this end, let $p \in X$. Then condition (iii) in the statement of the lemma is immediately ruled out, and so either (i) or (ii) must hold. If (i) holds, then there are geodesic arcs on either side of Γ which pass through p. Thus, by the uniqueness of geodesics, these arcs must coincide with each other, and consequently with Γ near p, as desired. If (ii) holds, then it follows from Lemma 2.1 (the maximum principle) that Γ again coincides with G_t near p, which completes the proof. □

Next we characterize convex spherical sets which contain a pair of antipodal points. A *lune* is a subset of $\mathbf{S}^2$ which is bounded by a pair of half great circles sharing the same end points.

Lemma 2.3. *Let X be a proper closed convex subset of $\mathbf{S}^2$ which contains a pair of antipodal points. Then X is a lune.*

Proof. Suppose $\pm p \in X$. Let $L \subset \mathbf{S}^2$ be the union of all half great circles ℓ_q which pass through $q \in X - \{\pm p\}$, and have their end points at $\pm p$. Then $X \subset L$. Further $L \subset X$, since for every $\ell_q \in L$, the subarcs pq and $(-p)q$ of ℓ_q both lie in X by the convexity assumption. So $L = X$, and it remains only to check that L is a lune. To see this, first note that L is closed since X is closed. Next, let $C \subset \mathbf{S}^2$ be the great circle which lies on the plane which is orthogonal to the line passing through $\pm p$, and define $\pi \colon \mathbf{S}^2 - \{\pm p\} \to C$ by $\pi(q) := \ell_q \cap C$. Note that $\pi(L) = C \cap L$. So $\pi(L)$ is an arc, or a closed connected proper subset of C, since L is closed, convex, and a proper subset of $\mathbf{S}^2$. Consequently L is a lune, since $L = \pi^{-1}(\pi(L)) \cup \{\pm p\}$. Indeed L is bounded by $\ell_r \cup \ell_s$ where r and s are the end points of $\pi(L)$. □

2.4. **The Beltrami map.** Let M and $\overline{M}$ be a pair of Riemannian surfaces, $f \colon M \to \overline{M}$ be a diffeomorphism, $\Gamma \subset M$ be a $\mathcal{C}^2$-regular simple curve, $p \in \Gamma$, N be a unit normal vector field along Γ, and k be the corresponding geodesic curvature. Further let $\overline{\Gamma} := f(\Gamma)$, $\overline{N}$ be the unit normal vector field along $\overline{\Gamma}$ which lies on the same side of $\overline{\Gamma}$ as $df(N)$, and $\overline{k}$ be the corresponding geodesic curvature of $\overline{\Gamma}$. We say that f *preserves the sign of geodesic curvature*, provided that $k(p) > 0$ (resp. < 0) if and only if $\overline{k}(\overline{p}) > 0$ (resp. $\overline{k}(\overline{p}) < 0$). By a *hemisphere* in this work we always mean a closed hemisphere unless noted otherwise. A pair of curves are said to have *contact of order n* at an intersection point provided that they admit regular parametrizations whose derivatives agree up to order n at that point. The maximum principle (Lemma 2.1) quickly yields:

Lemma 2.4. *Let M, $\overline{M}$ be Riemannian surfaces with nonnegative curvature, and $f\colon M \to \overline{M}$ be a diffeomorphism which preserves geodesics. Then f preserves the sign of geodesic curvature. In particular, for any open hemisphere H centered at a point $p \in \mathbf{S}^2$, the radial projection $\pi\colon H \to T_p\mathbf{S}^2 \simeq \mathbf{R}^2$ from the center of $\mathbf{S}^2$ preserves the sign of geodesic curvature.*

Proof. Note that $p \in \Gamma$ is an inflection, if, and only if, there passes a geodesic C through p which has contact of order 2 with Γ. Since f is a diffeomorphism, it preserves the order of contact, and so it will preserve inflections if it preserves geodesics. Further, by Lemma 2.1, $k(p) > 0$ (rep. < 0) if and only if $k(p) \neq 0$, and, near p, Γ lies in the same (resp. opposite) side of C where $N(p)$ points, which are all properties preserved by f. □

The above lemma essentially states that, as far as this work is concerned, local geometry of spherical curves is the same as that of planar curves, which will be a convenient observation utilized in the following pages. We should also point out that the last sentence in Lemma 2.4 is well known, e.g., see [46], although we are not aware of an explicit or concise proof. The projection from the center of the sphere is also called *gnomic projection* or *Beltrami map* [9].

Note 2.5. In addition to the Beltrami map, the stereographic projection $\pi\colon \mathbf{S}^2 - \{(0,0,1)\} \to \mathbf{R}^2$ is also useful for studying spherical curves, because it preserves their vertices. Indeed, a point of a curve in $\mathbf{R}^2$ or $\mathbf{S}^2$ is a vertex (local extremum of geodesic curvature), if, and only if, it lies locally on one side of its osculating circle at that point. It remains to note then that π preserves osculating circles, because it maps circles to circles and preserves the order of contact between curves. This observation implies that Segre's theorem, which is generalized by inequality (4), is itself a generalization of the classical theorem of Kneser which states that simple $\mathcal{C}^2$-regular closed curves $\Gamma \subset \mathbf{R}^2$ have (at least) 4 vertices (earlier, Mukhopadhyaya had obtained this result for convex curves). Indeed, we may assume, after a dilation and translation, that Γ intersects $\mathbf{S}^1$ in a pair of antipodal points $\pm p$. Then the spherical curve $\overline{\Gamma} := \pi^{-1}(\Gamma)$ also includes $\pm p$ and thus contains the origin in its convex hull. So, by Segre's theorem, $\overline{\Gamma}$ has 4 inflections, and therefore 4 vertices (there must be a local extremum of geodesic curvature k in between every pairs of zeros of k). So Γ must have 4 vertices as well.

3. Inflections of Curves Inscribed in Great Circles

Here we use the observations of the last section to prove Theorem 1.2 in the special case where Γ is simple, regular ($\mathcal{D}^+ = \mathcal{S} = 0$), and lies in a hemisphere. To this end we need the following basic lemma. By an *orientation* of a simple curve Γ, we mean an ordering of its points (which may be cyclical in case Γ is closed). If Γ is oriented, then for any pairs of points p, q of Γ, the oriented arc in Γ which begins at p and ends at q (if it exists) is denoted by pq. If Γ is $\mathcal{C}^1$, then this ordering corresponds to a unique choice of a unit tangent vector field along Γ, which will also be called an orientation of Γ. We say that a pair of oriented $\mathcal{C}^1$-regular curves are

cotangent at an intersection point p, if they are tangent at p and their orientations coincide at that point.

Lemma 3.1. *Let $\Gamma \subset \mathbf{S}^2$ be an oriented $\mathcal{C}^1$-regular simple arc which is $\mathcal{C}^2$-regular in its interior. Suppose that Γ lies on one side of an oriented great circle C, intersects C only at its end points p and q, and is cotangent to C at p. Further suppose that p is the initial point of Γ, and the length of the arc pq in C is at most π. Then the geodesic curvature k of interior of Γ changes sign at least once. Furthermore, if Γ is also cotangent to C at q, then k changes sign at least twice.*

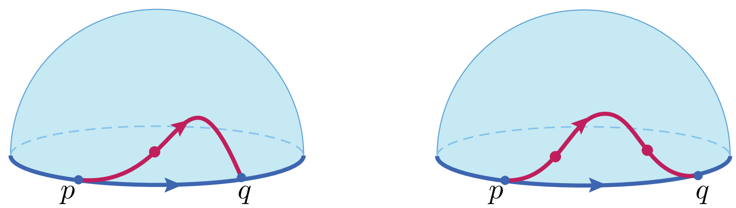


FIGURE 4.

Proof of Lemma 3.1. Let H be the hemisphere determined by C where Γ lies, and N be a unit normal vector field along Γ such that $N(p)$ points inside H. Note that if Γ is also cotangent to C at q, then $N(q)$ points inside H as well; see Figure 4. To see this, let $\overline{C} \subset H$ be the closed curve formed by joining Γ to the arc qp in C. Then $\overline{C}$ is simple, so it bounds a region $\overline{H} \subset H$ by the Jordan curve theorem. Further, $\overline{C}$ is $\mathcal{C}^1$-regular, so we may choose a unit normal vector field $\overline{N}$ along $\overline{C}$ which always points into $\overline{H}$. Now note that $N(p) = \overline{N}(p)$. Thus $N = \overline{N}$ all along Γ. In particular, $N(q) = \overline{N}(q)$, which shows that $N(q)$ points inside $\overline{H} \subset H$ as claimed.

Next let k be the geodesic curvature of the interior of Γ with respect to N. Then $k > 0$ arbitrarily close to p by Lemma 2.1. Furthermore, $k > 0$ arbitrarily close to q as well, if Γ is cotangent to C at q. So all we need is to show that $k < 0$ at some point in the interior of Γ. Suppose not. Then $\overline{C}$ is a convex curve by Lemma 2.2. But $\overline{C}$ contains antipodal points, because the length of the arc qp of C is at least π by assumption. So $\overline{C}$ must bound a lune by Lemma 2.3, which is impossible; because then Γ would have to be a geodesic, and therefore lie on C since it is tangent to C at p, but Γ may touch C only at its end points by assumption. □

The last lemma quickly yields the following observation, which corresponds to a theorem of Osserman for planar curves (see Note 3.4). We say that a continuous function $f\colon \mathbf{S}^1 \to \mathbf{R}$ *changes sign* n times, provided that there are cyclically arranged consecutive points $p_i \in \mathbf{S}^1$, $i \in \mathbf{Z}_n$, such that $f(p_i) \neq 0$, and $\operatorname{sign}(f(p_i)) = -\operatorname{sign}(f(p_{i+1}))$.

Theorem 3.2. *Let $\Gamma \subset \mathbf{S}^2$ be a $\mathcal{C}^2$-regular simple closed curve which is circumscribed in a great circle C, and has no geodesic subarcs. Suppose that $\Gamma \cap C$ includes n points not all contained in the interior of a semicircle of C. Then geodesic curvature of Γ changes sign at least $2n$ times. In particular, if Γ has only finitely many inflections, then at least $2n$ of them must be genuine.*

Proof. Let H be the hemisphere bounded by C where Γ lies, and N be the unit normal vector field along C which points into H. Orient C by choosing a unit tangent vector field T along it so that (T, N) lies in a fixed orientation class of $\mathbf{S}^2$ (e.g., assume that $\det(T, N, n) > 0$ where n is the outward unit normal of $\mathbf{S}^2$). Now let p_i, $i \in \mathbf{Z}_n$, be a cyclical ordering of the given points in $\Gamma \cap C$ with respect to the orientation of C. We need to show that the curvature changes sign (at least) twice on each (oriented) subarc Γ_i with initial point p_i and final point p_{i+1}.

To this end first note that, since Γ has no geodesic subarcs by assumption, there is a point $q_i \in \Gamma_i$ which does not lie on C. Let $\overline{\Gamma}_i$ be the largest subarc of Γ_i which contains q_i, and whose interior is disjoint from C. Then $\overline{\Gamma}_i$ will meet C only at its end points $\overline{p}_i$ and $\overline{p}_{i+1}$. Further, since Γ is a simple closed curve, $\overline{p}_i$ and $\overline{p}_{i+1}$ must lie in the arc $p_i p_{i+1} \subset C$ which by assumption has length at most π. Thus $\overline{p}_i\overline{p}_{i+1} \subset C$ will also have length at most π.

Now we claim that $\overline{\Gamma}_i$ is cotangent to C at its end points, which will complete the proof by Lemma 3.1. This follows from the assumption that Γ is simple. Indeed, let $\widetilde{H} \subset H$ be the region bounded by Γ, and $\widetilde{N}$ be the unit normal vector field of Γ which points into $\widetilde{H}$. Next let $\widetilde{T}$ be the unit tangent vector field of Γ such that $(\widetilde{T}, \widetilde{N})$ lies in the preferred orientation of class of $\mathbf{S}^2$, which we fixed earlier. Then $N(p) = \widetilde{N}(p)$ at every point $p \in \Gamma \cap C$, which in turn yields that $T(p) = \widetilde{T}(p)$ as desired. □

Finally we are ready to prove the main result of this section, which follows from the last theorem via some basic convexity theory. For any set $X \subset \mathbf{R}^n$, let $\operatorname{conv}(X)$ denote the *convex hull* of X, i.e., the intersection of all closed half-spaces which contain X.

Corollary 3.3. *Theorem 1.2 holds if Γ is simple, regular ($\mathcal{D}^+ = \mathcal{S} = 0$), and lies in a hemisphere.*

Proof. First note that (5) holds trivially here, since if Γ lies in a hemisphere and is symmetric, then it must be a great circle (and thus have infinitely many inflections). So we just need to consider (3) and (4).

Since $o \in \operatorname{conv}(\Gamma)$, o lies in a convex polytope P with vertices on Γ, by the theorem of Carathéodory [37]. Next note that, since Γ lies in a hemisphere it lies on one side of a plane $\Pi \subset \mathbf{R}^3$ which passes through o. In particular, o lies on the boundary of $\operatorname{conv}(\Gamma)$, and therefore on the boundary of P. More specifically, o lies in $F := \Pi \cap P$, which is either an edge or a face of P.

If F is an edge, then the vertices of F are antipodal points of $\mathbf{S}^2$, since $o \in F$. Thus if $C := \Pi \cap \mathbf{S}^2$ denotes the great circle determined by Π, then $\Gamma \cap C$ consists of 2 points which are not all contained in an open semicircle of C. So, by Theorem 3.2, $\mathcal{I} \geq 4$, which yields (4). Further, we already have $\mathcal{D} \geq 1$ due to the existence of a pair of antipodal points on Γ. So $2\mathcal{D} + \mathcal{I} \geq 6$ which yields (3).

If, on the other hand, F is a face, then it must have at least 3 vertices. Since $o \in F$, not all these vertices can lie in an open semicircle of C. Thus Theorem 3.2 yields that $\mathcal{I} \geq 6$, which completes the proofs of (3) and (4). □

Note 3.4. Theorem 3.2 extends a result of Jackson [26, Thm. 7.1], see also [33, 27], which shows that if a simple $\mathcal{C}^2$-regular closed curve $\Gamma \subset \mathbf{R}^2$ touches its circumscribing circle C in n points, then it must have $2n$ vertices (where "vertex" here means a local extremum of geodesic curvature, and C is the smallest circle which contains Γ). This follows by applying the stereographic projection $\pi\colon \mathbf{S}^2 - \{(0,0,1)\} \to \mathbf{R}^2$ along the lines discussed in Note 2.5. Indeed, after a rigid motion and dilation, we may assume that $C = \mathbf{S}^1$. Then $\overline{\Gamma} := \pi^{-1}(\Gamma)$ lies in the Southern hemisphere, and touches the equator C at $n \geq 2$ points not all of which are contained in an open semicircle. So, by Theorem 3.2, $\overline{\Gamma}$ has (at least) $2n$ inflections, and therefore $2n$ vertices. Hence Γ has $2n$ vertices as well, since π preserves vertices.

4. Curve Shortening Flow: A New Proof of Segre's Theorem

Here we apply the observations of the last section, as summed up in Corollary 3.3, to prove inequality (3) in the special case where Γ is regular and simple ($\mathcal{D}^+ = \mathcal{S} = 0$). Without much additional effort, we will also obtain in this section proofs of the inequalities (4) and (5) in the simple regular case. These constitute new proofs of the theorems of Segre and Möbius, which were mentioned in the introduction. The main strategy here is to use the curve shortening flow [7, 22, 24, 3] to deform Γ through simple $\mathcal{C}^2$-regular curves, without increasing the number of inflections or antipodal pairs of points of Γ, until Γ just lies in a hemisphere, in which case results of the previous section may be applied. The curve shortening flow is ideally suited to this task due to a number of well known remarkable properties which we summarize below for the reader's convenience. A $\mathcal{C}^{2,1}$ curve is one which has a $\mathcal{C}^{2,1}$ parametrization (i.e., a $\mathcal{C}^2$ parametrization with Lipschitz continuous second derivative). A simple closed curve $\Gamma \subset \mathbf{S}^2$ is said to *bisect* $\mathbf{S}^2$ if the components of $\mathbf{S}^2 - \Gamma$ have equal area.

Lemma 4.1. *Let $\Gamma \subset \mathbf{S}^2$ be a simple $\mathcal{C}^{2,1}$-regular closed curve. There exists a continuous one parameter family $\Gamma_t \subset \mathbf{S}^2$, $0 \leq t < t_1$, of simple $\mathcal{C}^2$-regular closed curves with $\Gamma_0 = \Gamma$ such that*

(i) *Γ_t converges to a point, as $t \to t_1$, unless Γ bisects $\mathbf{S}^2$.*
(ii) *The number of inflections of Γ_t is nonincreasing; furthermore, if Γ_0 has only finitely many inflections, the number of* genuine *inflections of Γ_t is nonincreasing as well.*
(iii) *The number of intersections of Γ_t with $-\Gamma_t$ is nonincreasing.*

Proof. Let $\gamma_0 \colon \mathbf{S}^1 \to \Gamma$ be a $\mathcal{C}^{2,1}$ regular parametrization, and γ_t be the evolution of γ_0 according to the curve shortening flow, i.e., the solution to the equation

$$\frac{\partial}{\partial t}\gamma_t(s) = k_t(s)N_t(s),$$

where N_t is a continuous choice of unit normal vector fields along γ_t, and k_t is the corresponding geodesic curvature. By Grayson's theorem [22] $\Gamma_t := \gamma_t(\mathbf{S}^1)$ is a simple $\mathcal{C}^\infty$-regular closed curve, for some interval $0 < t < t_1$, which converges either to a closed geodesic or a point as $t \to t_1$. Let Ω_t be a continuous choice of regions in $\mathbf{S}^2$ bounded by Γ_t, $A(t)$ be the area of Ω_t, and suppose that N_t always

points into Ω_t. After possibly replacing Ω_t with its complements, we may assume that $A(0) \leq 2\pi$. In particular, if Γ_t does not bisect $\mathbf{S}^2$, we have $A(0) < 2\pi$.

It follows from the Gauss-Bonnet theorem that $A'(t) = A(t) - 2\pi$ [3]. Thus $A(t)$ is decreasing, and so it cannot approach 2π. Consequently Γ_t cannot converge to a great circle, and must therefore converge to a point by Grayson's theorem. This settles item (i). For item (ii) see [3] or [22, Cor. 2.5] which establish that the number of inflections never increases; furthermore, see [22, Lem. 1.5], where it is shown that for $t > 0$ the only times when k_t and k_t' both vanish at the same point are when two or more inflections merge. In particular, non-genuine inflections exist only at isolated times before disappearing, and the non-genuine inflections of Γ_0 immediately disappear as well. Finally, item (iii) follows from a theorem of Angenent [2, p. 179] who showed that the number of intersections of a pair of curves under the curve shortening flow is nonincreasing. □

To apply the curve shortening flow we need the initial curve to be $\mathcal{C}^{2,1}$ according to Lemma 4.1; however, in Theorem 1.2 we are assuming only that the given curve is $\mathcal{C}^2$. To resolve this discrepancy we need the following basic approximation result:

Lemma 4.2. *Let $f\colon \mathbf{S}^1 \to \mathbf{R}$ be a continuous function with a finite number of zeros. Then for any given $\epsilon > 0$ there exists a $\mathcal{C}^\infty$ function $\widetilde{f}\colon \mathbf{S}^1 \to \mathbf{R}$ such that $\|\widetilde{f} - f\| \leq \epsilon$, and $\widetilde{f}$ has no more zeros than f has.*

Proof. Let $t_i \in \mathbf{S}^1 \simeq \mathbf{R}/2\pi$, $i = 1, \dots, n$, be the zeros of f, and choose $\delta > 0$ sufficiently small so that $f \neq 0$ on $[t_i - \delta, t_i) \cup (t_i, t_i + \delta]$. Next let $\overline{f}\colon \mathbf{S}^1 \to \mathbf{R}$ be the function which is obtained from f by replacing the graph of f over each of the intervals $[t_i - \delta, t_i + \delta]$ with the line segment joining the points $(t_i - \delta, f(t_i - \delta))$ and $(t_i + \delta, f(t_i + \delta))$. Choosing δ sufficiently small, we may assume that $\|f\| \leq \epsilon/4$ on $[t_i - \delta, t_i + \delta]$, which yields that $\|\overline{f}\| \leq \epsilon/4$ on $[t_i - \delta, t_i + \delta]$ as well. So it follows that $\|\overline{f} - f\| \leq \epsilon/2$. Also note that $\overline{f}$ has no more zeros than f has, $\overline{f}$ is linear on $[t_i - \delta, t_i + \delta]$, and $\overline{f} \equiv f \neq 0$ outside these neighborhoods.

For $\lambda > 0$, let $\theta_\lambda\colon \mathbf{R} \to \mathbf{R}$ be a $\mathcal{C}^\infty$ nonnegative function such that $\int_{\mathbf{R}} \theta_\lambda = 1$, and $\theta_\lambda \equiv 0$ outside $(-\lambda, \lambda)$. Also let $\widetilde{f}_\lambda := \overline{f} * \theta_\lambda$ be the convolution of $\overline{f}$ with θ_λ. Choosing λ sufficiently small, we may assume that $\|\widetilde{f}_\lambda - \overline{f}\| \leq \epsilon/2$. Thus,

$$\|\widetilde{f}_\lambda - f\| \leq \|\widetilde{f}_\lambda - \overline{f}\| + \|\overline{f} - f\| = \epsilon.$$

Further, $\widetilde{f}_\lambda \equiv \overline{f}$ on each interval $I_i := [t_i - \delta + \lambda, t_i + \delta - \lambda]$, since $\overline{f}$ is linear on $[t_i - \delta, t_i + \delta]$. So $\widetilde{f}_\lambda$ has at most one zero in each I_i. Also note that $\overline{f} \neq 0$ outside the interiors of I_i (which is a compact region). Thus, since $\widetilde{f}_\lambda \to \overline{f}$, as $\lambda \to 0$, we may choose λ so small that $\widetilde{f}_\lambda \neq 0$ outside I_i. Then $\widetilde{f}_\lambda$ is the desired function. □

The last lemma leads to the next observation. Here $\|\cdot\|_{\mathcal{C}^2}$ denotes the standard (supremum) norm in the space of $\mathcal{C}^2$ maps $f\colon I \to \mathbf{R}^n$.

Proposition 4.3. *Let $\gamma\colon \mathbf{S}^1 \to \mathbf{S}^2$ be a regular $\mathcal{C}^2$ closed curve with finitely many inflections. Then for every $\epsilon > 0$ there is a regular $\mathcal{C}^\infty$ closed curve $\widetilde{\gamma}\colon \mathbf{S}^1 \to \mathbf{S}^2$ which has no more inflections than γ has, and $\|\widetilde{\gamma} - \gamma\|_{\mathcal{C}^2} \leq \epsilon$.*

Proof. Suppose that γ has constant speed, let k be the geodesic curvature of γ, and $\overline{k}\colon \mathbf{S}^1 \to \mathbf{R}$ be the smoothing of k given by Lemma 4.2. We may identify $\overline{k}$ with a 2π-periodic function on $\mathbf{R}$, and assume that $\overline{k}(0) \neq 0$. Now let $\overline{\gamma}\colon \mathbf{R} \to \mathbf{S}^2$ be the curve with geodesic curvature $\overline{k}$ such that $\overline{\gamma}(0) = \gamma(0)$ and $\overline{\gamma}'(0) = \gamma'(0)$. Then $\overline{\gamma}$ is $\mathcal{C}^\infty$ and, choosing $\overline{k}$ sufficiently close to k, we may assume that $\|\overline{\gamma} - \gamma\|_{\mathcal{C}^2} \leq \epsilon$ on $[-2\pi, 4\pi]$. This follows from basic ODE theory which guarantees the existence and uniqueness of $\overline{\gamma}$ together with its continuous dependence on $\overline{k}$. Indeed, $\overline{\gamma} \to \gamma$ with respect to the $\mathcal{C}^2$ norm as $\overline{k} \to k$.

Next we are going to "close up" $\overline{\gamma}$ by a perturbation as follows. Assume that δ is so small that γ has no inflections on $[-\delta, \delta]$. Let $\phi\colon [0, 2\pi+\delta] \to \mathbf{R}$ be any $\mathcal{C}^\infty$ nonincreasing function such that $\phi \equiv 1$ on $[0, 2\pi-\delta]$ and $\phi \equiv 0$ on $[2\pi, 2\pi+\delta]$. Now define $\widetilde{\gamma}\colon [0, 2\pi+\delta] \to \mathbf{R}^3$ by

$$\widetilde{\gamma}(t) := \phi(t)\overline{\gamma}(t) + \big(1-\phi(t)\big)\overline{\gamma}(t-2\pi).$$

Then $\widetilde{\gamma}(t) = \overline{\gamma}(t) = \widetilde{\gamma}(t+2\pi)$ on $[0,\delta]$. So setting $\widetilde{\gamma}(t+2\pi) := \widetilde{\gamma}(t)$, and then replacing $\widetilde{\gamma}$ with $\widetilde{\gamma}/\|\widetilde{\gamma}\|$, we obtain a regular $\mathcal{C}^\infty$ closed curve $\widetilde{\gamma}\colon \mathbf{S}^1 \simeq \mathbf{R}/2\pi \to \mathbf{S}^2$. Since $\overline{\gamma}$ may be chosen arbitrarily $\mathcal{C}^2$-close to γ, it follows that $\widetilde{\gamma}$ may be arbitrarily $\mathcal{C}^2$-close to γ. In particular, since γ has no inflections on $[-\delta,\delta]$, then we may assume that $\widetilde{\gamma}$ has no inflections on that interval either. Thus $\widetilde{\gamma}$ has no more inflections than γ has. $\square$

Finally we need the following basic result which will allow us to perturb curves locally without increasing the number of their isolated inflections. This lemma, which will be invoked frequently in the following pages, applies to all curves which may be represented locally as the graph of a function $f\colon [a,b] \to \mathbf{R}$ (e.g., all curves which are $\mathcal{C}^1$ or convex). Note that the graph of f has an inflection at $(x, f(x))$, if, and only if, $f''(x) = 0$. We say that a pair of functions f_1, $f_2\colon [a,b] \to \mathbf{R}$ have the *same regularity properties* provided that f_1 is $\mathcal{C}^k$ on a subset $U \subset [a,b]$, if, and only if, f_2 is $\mathcal{C}^k$ on U.

Lemma 4.4. *Let $f\colon [-a,a] \to \mathbf{R}$ be a continuous function. Suppose that f is $\mathcal{C}^2$ on $[-a,0) \cup (0,a]$, and $f'' \neq 0$ except possibly at 0. Then for any $\epsilon > 0$, there is a function $\widetilde{f}\colon [-a,a] \to \mathbf{R}$ with the same regularity properties as f, such that $\widetilde{f} \geq f$, $\widetilde{f}(0) > f(0)$, $\widetilde{f} \equiv f$ near $\pm a$, $\|\widetilde{f} - f\| \leq \epsilon$, and in case f is $\mathcal{C}^2$, $\|\widetilde{f} - f\|_{\mathcal{C}^2} \leq \epsilon$. Finally, f_ϵ'' has no more zeros than f'' has.*

Proof. Let $\phi\colon [-a,a] \to \mathbf{R}$ be a $\mathcal{C}^\infty$ nonnegative function such that $\phi \equiv 0$ outside $(-a/2, a/2)$ and $\phi \equiv 1$ on $[-a/4, a/4]$. Set $\widetilde{f_\delta} := f + \delta\phi$. It is easy to see that, for sufficiently small $\delta > 0$, $\widetilde{f_\delta}$ has all the required properties except possibly the last one. To see that the last requirement is met as well, note that $\widetilde{f_\delta}'' \equiv f''$ outside of $A := [-a/2, -a/4] \cup [a/4, a/2]$. Thus it suffices to check that $\widetilde{f_\delta}'' \neq 0$ on A. This is indeed the case since as $\delta \to 0$, $\widetilde{f_\delta} \to f$ with respect to the $\mathcal{C}^2$ norm on A, and $f'' \neq 0$ on A, by assumption. $\square$

We are now ready to prove the main result of this section:

Proposition 4.5. *Theorem 1.2 holds when* Γ *is simple and regular* ($\mathcal{D}^+ = \mathcal{S} = 0$).

One step in the proof of the above result worth being highlighted is the following observation which shows that Möbius's theorem follows from Segre's result:

Lemma 4.6. *Inequality* (4) *implies* (5) *when* $\mathcal{D}^+ = \mathcal{S} = 0$.

Proof. Suppose that $\Gamma = -\Gamma$. We may assume that Γ has only finitely many inflections. Then the number of genuine inflections $\mathcal{I} \geq 4$ by (4). We need to show that $\mathcal{I} \geq 6$. Since $\mathcal{I}$ must be even, it suffices then to show that $\mathcal{I} > 4$. Suppose, towards a contradiction, that $\mathcal{I} = 4$. Then the sign of the four arcs of Γ in between these inflections alternate as we go around Γ. In particular, we end up with a pair of opposite arcs with the same sign. This is the desired contradiction because antipodal reflection in $\mathbf{S}^2$ switches the sign of geodesic curvature, i.e., $k(-p) = -k(p)$ for every point $p \in \Gamma$. $\square$

Next we complete the rest of the proof:

Proof of Proposition 4.5. By Lemma 4.6, it remains to establish the inequalities (3) and (4). First note that if $\mathcal{D}^+ = \mathcal{S} = 0$ in Theorem 1.2, then Γ is a $\mathcal{C}^2$-regular simple closed curve. Now if Γ lies in a hemisphere, then Theorem 3.2 completes the proof. So we may assume that the origin o lies in the interior of $\mathrm{conv}(\Gamma)$. Then, by Proposition 4.3, we may assume after a perturbation that Γ is $\mathcal{C}^{2,1}$, and thus curve shortening results of Lemma 4.1 may be applied to Γ. Also note that Γ remains simple after this perturbation, since it is arbitrarily small with respect to the $\mathcal{C}^1$-topology. Now set $\Gamma_0 := \Gamma$ and let Γ_t be the corresponding flow as in Lemma 4.1. Since, by Lemma 4.1, the number of genuine inflections and antipodal pairs of points of Γ_t do not increase with time (not to mention that Γ_t remains simple), it suffices to show that Γ_t will satisfy (3) and (4) at some future time.

To take advantage of the last observation note that, by Lemma 4.4, we may assume that Γ_0 does not bisect $\mathbf{S}^2$ after a perturbation. Indeed this perturbation may be confined to an arbitrarily small neighborhood of a point p of Γ_0 which is neither an inflection nor an element of an antipodal pair of points of Γ. More specifically, a small neighborhood U of such a point may be identified with a ball in $\mathbf{R}^2$ via Lemma 2.4. Then an arc of Γ_0 containing p may be identified with the graph of a function $f\colon [-a, a] \to \mathbf{R}$. Then Lemma 4.4 yields the desired perturbation of the graph of f, and consequently of Γ_0.

So we may assume that Γ_t converges to a point as $t \to t_1$, by Lemma 4.1. In particular, Γ_t eventually lies in a hemisphere. Set

$$t_0 := \inf \big\{ t \in [0, t_1) \mid \Gamma_t \text{ lies in a hemisphere} \big\}. \tag{9}$$

Then Γ_{t_0} lies in a hemisphere (because according to our definition hemispheres are closed). Furthermore, $o \in \mathrm{conv}(\Gamma_{t_0})$. If not, then there exists a plane passing through o which is disjoint from $\mathrm{conv}(\Gamma_{t_0})$. Consequently, Γ_{t_0} lies in the interior of a hemisphere H. This yields that $\Gamma_{t_0-\epsilon}$ must have lied in the interior of H as well, for some $\epsilon > 0$, which contradicts (9). Thus Γ_{t_0} satisfies (3) and (4) by Corollary 3.3. $\square$

Note 4.7. The curve shortening flow has been used by Angenent [3] to prove Segre's theorem in the special case where Γ bisects $\mathbf{S}^2$ (the case known as Arnold's "tennis ball theorem"). In that setting Γ flows to a great circle, and the Sturm theory is used to show that perturbations of great circles have at least 4 inflections, if they contain the origin in their convex hulls. In comparison, our proof here is more elementary (as well as more general), since we avoid using Sturm theory by an initial perturbation of the curve, which ensures that it will flow into a hemisphere, where Theorem 3.2 may be applied.

5. Removal of Singularities

Having proved Theorem 1.2 in the simple regular case ($\mathcal{D}^+ = \mathcal{S} = 0$), we now turn to the general case. To this end we need a desingularization result developed in this section, which shows that the singularities of a $\mathcal{C}^2$ curve $\Gamma \subset \mathbf{S}^2$ may be eliminated without increasing the sum

$$\Sigma^+(\Gamma) := 2(\mathcal{D}^+ + \mathcal{S}) + \mathcal{I}. \tag{10}$$

Note that as far as proving Theorem 1.2 is concerned, we may assume that all multiple points of Γ are double points, for if $\gamma^{-1}(p)$ has more than 2 elements for some $p \in \Gamma$, then $\mathcal{D}^+ \geq 3$, which yields that $\Sigma^+(\Gamma) \geq 6$ and so there would be nothing left to prove. A multiple point p is called a *multiple singularity* if at least two of the points in the preimage $\gamma^{-1}(p)$ are singular points of γ. If Γ has a multiple singularity, then $(\mathcal{D}^+ + \mathcal{S}) \geq 3$ which again yields $\Sigma^+(\Gamma) \geq 6$. Thus, as far as proving theorem 1.2 is concerned, we may assume that Γ has no multiple singularities.

Theorem 5.1. *Let M be a Riemannian surface of constant nonnegative curvature, $\Gamma \subset M$ be a $\mathcal{C}^2$ arc with finitely many singularities and inflections, and p be a singular point of Γ which is either a simple point, or a double point, but is not a multiple singularity. Then for every open neighborhood U of p in M, which does not contain any other singularities, there is a $\mathcal{C}^2$ arc $\widetilde{\Gamma}$ such that $\widetilde{\Gamma} = \Gamma$ outside U, $\widetilde{\Gamma}$ is $\mathcal{C}^2$-regular inside U, and $\Sigma^+(\widetilde{\Gamma}) \leq \Sigma^+(\Gamma)$.*

The rest of this section will be devoted to proving the above theorem, which constitutes the most technical part of the paper. This will unfold in two parts: first we consider the case of simple singularities, and then that of singularities which are double points (but are not multiple singularities). Note that since the theorem is essentially local, we may assume that $M = \mathbf{R}^2$ via Lemma 2.4. Indeed, assuming that the closure of U is a sufficiently small metric ball centered at p, we may identify it (isometrically) with a ball in $\mathbf{R}^2$, in case the curvature of M is zero; otherwise, we may isometrically map the closure of U to a ball in an open hemisphere of $\mathbf{S}^2$ after a uniform rescaling of the metric, and then map it to $\mathbf{R}^2$ via the Beltrami projection described in Section 2.4.

5.1. **Simple singularities.** Here we suppose that p in Theorem 5.1 is a simple point. Then, after replacing Γ with a subarc, we may assume that Γ is a simple arc without any inflections and singularities except possibly at p. Further we may assume that Γ lies in a ball B, p coincides with the center o of B, and Γ intersects

∂B only at its end points. Then Γ will be called a *double spiral*. For easy reference, let us record that:

Definition 5.2. A simple arc $\Gamma \subset \mathbf{R}^2$ is a *spiral*, if it is $\mathcal{C}^2$-regular in the complement of a boundary point o, called the *vortex* of Γ, has no inflections except possibly at o, lies in a ball B centered at o, and touches ∂B only at its other end point. Further, Γ is a *double spiral* if it is the union of two spirals, called the *arms* of Γ, which lie in the same ball B, and intersect only at their common vortex o in the center of B.

Note that we do not assume that a spiral Γ has any regularity at its vortex o. In particular, the tangent lines of Γ may not converge to a limit at o (Γ may rotate endlessly about o). Recall that the closure of each component of $B - \Gamma$ is called a side S of Γ. We say S is *proper* provided that there exists no line segment in S passing through o. If Γ has no proper sides, then it must coincide with a line segment in a neighborhood of o, which implies that the arms of Γ must contain inflections—a contradiction. So *every double spiral has at least one proper side.*

Double spirals may be categorized into three types as follows. By a *principal normal* vector field of a double spiral Γ we mean a unit normal vector field along $\Gamma - \{o\}$ such that the geodesic curvature of Γ with respect to N is always positive. Since by assumption $\Gamma - \{o\}$ has no inflections, N always exists. If N only points into one side S of Γ, then Γ is called either *convex* or *concave* according to whether or not S is proper. Otherwise (if N changes sides), we say that Γ is *semiconvex*. Examples of these three types of double spirals are depicted in Figure 5. As this

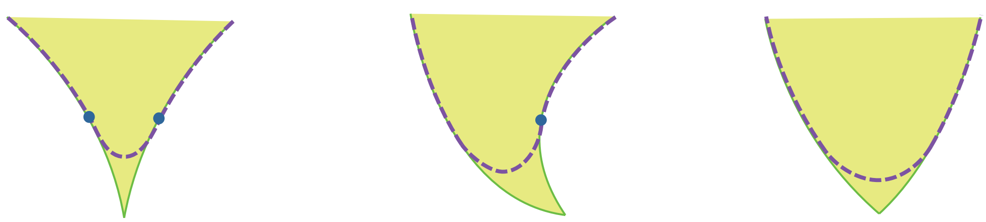


FIGURE 5.

figure suggests, in each case we may remove the singularity which may exist at the vortex o, by a perturbation of Γ near o, to obtain a simple $\mathcal{C}^2$-regular curve with at most two inflections. This is the content of the next result:

Proposition 5.3. *Let Γ be a double spiral with vortex o, and S be a proper side of Γ. Then there exists a simple $\mathcal{C}^2$-regular arc $\widetilde{\Gamma} \subset S - \{o\}$ which coincides with Γ near its end points, and has at most two inflections. Furthermore, if Γ is semiconvex or convex, then we may require that $\widetilde{\Gamma}$ have precisely 1 or 0 inflections respectively.*

The case of simple singularities in Theorem 5.1 is an immediate consequence of the above proposition. The proof of this proposition will be divided into three parts:

5.1.1. *The convex case.* First we prove Proposition 5.3 in the case where Γ is a convex double spiral. In general a simple curve Γ in $\mathbf{R}^2$ or $\mathbf{S}^2$ is called *convex* if

it lies on the boundary of a convex body (convex set with interior points), and the following lemma confirms that this is indeed the case with convex double spirals.

Lemma 5.4. *A convex double spiral is a convex curve.*

Proof. Let S be a proper side associated with a convex double spiral Γ. Then $\Gamma \subset \partial S$ and so it suffices to show that S is convex. To this end note that ∂S satisfies the hypothesis of Lemma 2.2, which has a direct analogue in $\mathbf{R}^2$ via Lemma 2.4. □

So a convex double spiral Γ may be represented as the graph of a convex function near its vortex o. Now, by Lemma 4.4, we may perturb Γ near o to obtain a convex double spiral $\overline{\Gamma}$ which lies in the convex side S of Γ, coincides with Γ near its end points, and its vortex $\overline{o}$ lies in the interior of S. Then there is a ball $\overline{B}$ centered at $\overline{o}$ which is disjoint from Γ. To complete the proof of the convex case of Proposition 5.3, it suffices then to show that we may perturb $\overline{\Gamma}$ within $\overline{B}$ to obtain a $\mathcal{C}^2$-regular curve without inflections, which is precisely the content of the next lemma:

Lemma 5.5. *Let Γ be a convex double spiral, and $U \subset \mathbf{R}^2$ be an open neighborhood of Γ. Then there exists a $\mathcal{C}^2$-regular simple arc $\widetilde{\Gamma} \subset U$ which coincides with Γ near its end points and has no inflections.*

Proof. After a rigid motion, we may identify a neighborhood of the vortex o of Γ with the graph of a convex function $f\colon [-a,a] \to \mathbf{R}$. The desired curve $\widetilde{\Gamma}$ is then obtained by replacing the graph of f with that of the function $\widetilde{f}\colon [-a,a] \to \mathbf{R}$ constructed as follows.

As in the proof of Lemma 4.2, let $\theta_\lambda\colon \mathbf{R} \to \mathbf{R}$ be a $\mathcal{C}^\infty$ nonnegative function such that $\int_{\mathbf{R}} \theta_\lambda = 1$, and $\theta_\lambda \equiv 0$ outside $(-\lambda, \lambda)$. Next let $f_\lambda\colon [-a+\lambda, a-\lambda] \to \mathbf{R}$ be the convolution of f with θ_λ, i.e., set

$$f_\lambda(x) := \int_{-a}^{a} f(x-y)\theta_\lambda(y)dy.$$

Then, $f_\lambda''(x) = \int_{-a}^{a} f''(x-y)\theta_\lambda(y)dy$. So, since $f'' > 0$ almost everywhere, it follows that $f_\lambda'' > 0$. Now let $\phi\colon [-a,a] \to \mathbf{R}$ be a $\mathcal{C}^\infty$ nonnegative function such that $\phi \equiv 1$ on $[-a/4, a/4]$, and $\phi \equiv 0$ on $[-a,-a/2] \cup [a/2, a]$. Assuming $\lambda < a/2$, and extending f_λ to $[-a,a]$ arbitrarily, we may define $\widetilde{f}_\lambda\colon [-a,a] \to \mathbf{R}$ by

$$\widetilde{f}_\lambda(x) := \big(1-\phi(x)\big)f(x) + \phi(x)f_\lambda(x).$$

We claim that, for sufficiently small λ, $\widetilde{f}_\lambda$ is the desired function.

To check this claim note that as $\lambda \to 0$, $f_\lambda \to f$, and thus $\widetilde{f}_\lambda \to f$. So, choosing λ sufficiently small, we may assume that $\|\widetilde{f}_\lambda - f\|$ is small enough so that $\widetilde{\Gamma}_\lambda \subset U$, where $\widetilde{\Gamma}_\lambda$ is the curve which is obtained from Γ by replacing the graph of f with that of $\widetilde{f}_\lambda$. Further note that $\widetilde{f}_\lambda'' > 0$ on $[-a,-a/2] \cup [-a/4, a/4] \cup [a/2, a]$, since f'', $f_\lambda'' > 0$ on this region. Next we show that for λ sufficiently small, $\widetilde{f}_\lambda'' > 0$ on $A := [-a/2, -a/4] \cup [a/4, a/2]$ as well, and thus $\widetilde{\Gamma}_\lambda$ is free of inflections, which would complete the proof. To see this note that on A

$$\widetilde{f}_\lambda'' = f'' + \phi''(f_\lambda - f) + 2\phi'(f_\lambda - f)' + \phi\,(f_\lambda - f)''.$$

Further, $f_\lambda \to f$ with respect to the $\mathcal{C}^2$ norm on A. Thus, since ϕ is independent of λ, $\widetilde{f}''_\lambda \to f''$ on A. So $\widetilde{f}''_\lambda > 0$ on A for small λ, because $f'' > 0$ on A. □

Note 5.6. Lemma 5.5 shows that an isolated inflection which lies in an open neighborhood where the curvature does not change sign may be removed by a local $\mathcal{C}^2$-small perturbation.

5.1.2. *The $\mathcal{C}^1$-regular case.* Next we show that Proposition 5.3 holds when Γ admits a regular $\mathcal{C}^1$ parametrization, or equivalently, Γ has a well defined tangent line which varies continuously along the curve. To this end first note that:

Lemma 5.7. *If Γ is a $\mathcal{C}^1$-regular double spiral, then Γ is either convex or semiconvex.*

Proof. We may suppose that the principal normal vector field N of Γ always points to one side of Γ, for otherwise Γ is semiconvex and there is nothing to prove. Let S be the side into which N points, and L be a line segment which passes through the vertex o of Γ. Then it follows from Lemma 2.1 that L lies outside of S near o. Now Lemma 2.2 shows that S is convex, which completes the proof (recall that Lemma 2.2 also holds in $\mathbf{R}^2$ via the Beltrami map or Lemma 2.4). □

Now, since the convex case has already been treated, we may assume that Γ is semiconvex. In this case we once again use Lemma 4.4 to perturb Γ inside the given proper side S, so that the vertex enters the interior of S, just as we had done in Section 5.1.1. Then all that remains is to show:

Lemma 5.8. *Let Γ be a $\mathcal{C}^1$-regular semiconvex double spiral, and $U \subset \mathbf{R}^2$ be an open neighborhood of Γ. Then there exists a $\mathcal{C}^2$-regular simple arc $\widetilde{\Gamma} \subset U$ which coincides with Γ near its end points, and has precisely one inflection.*

Proof. Much like the beginning of the proof of Lemma 5.5, we may identify a neighborhood of the vertex o of Γ with the graph of a function $f\colon [a, b] \to \mathbf{R}$. Replacing the graph of f with that of an appropriate function $\widetilde{f}\colon [a, b] \to \mathbf{R}$ then yields the desired curve $\widetilde{\Gamma}$. We may assume, after a rigid motion, that $f(0) = f'(0) = 0$, $f'' > 0$ on $(0, b]$, and $f'' < 0$ on $[a, 0)$. Then we construct $\widetilde{f}$ as follows.

Consider the cubic function $g(x) := \lambda(x^3 + x)$, where $\lambda > 0$. Choosing λ sufficiently small, we can make sure that $g(b) < f(b)$, and $g(a) > f(a)$. On the other hand, since $g(0) = f(0)$, while $g'(0) = \lambda > 0 = f'(0)$, it follows that $g > f$ on some interval $(0, b')$. Let $0 < b' < b$ be the largest number so that $g > f$ on $(0, b')$. Similarly, there exists $a < a' < 0$ such that $(a', 0)$ is the largest open interval ending at 0 where $g < f$. See Figure 6.

Now define $\overline{f}\colon [a, b] \to \mathbf{R}$ by setting $\overline{f} := g$ on $[a', b']$ and $\overline{f} := f$ otherwise. Then $\overline{f}$ is convex on $[0, b]$ and concave on $[a, 0]$, because the maximum (resp. minimum) of two convex (resp. concave) functions is convex (resp. concave). Note that $\overline{f}$ is $\mathcal{C}^2$ everywhere except near a' and b', and the only place where $\overline{f}''$ vanishes is at 0. Now we may smooth $\overline{f}$ near a' and b' by invoking Lemma 5.5 in small neighborhoods of these points to obtain the desired function $\widetilde{f}$. Finally note that, after replacing Γ

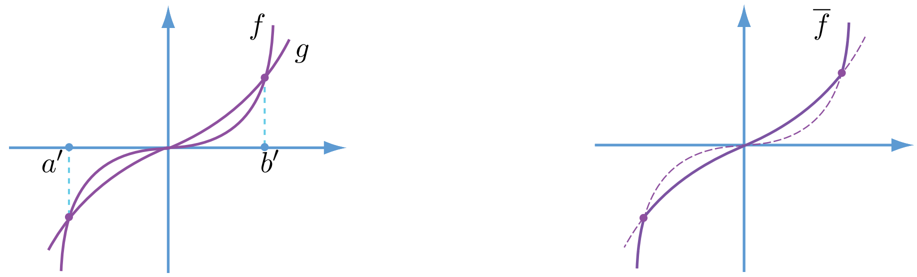


FIGURE 6.

with a subarc, we may assume that U is the interior of an open ball centered at o, in which case it is clear that the graph of $\widetilde{f}$ will lie within U as well. □

5.1.3. *Other cases.* It remains now to prove Proposition 5.3 in the case where Γ is neither convex nor $\mathcal{C}^1$-regular. To this end first note that

Lemma 5.9. *Let Γ be a spiral with vertex o, p be an interior point of Γ, and ℓ be the tangent line of Γ at p. Then the subarc po of Γ intersects ℓ only at p.*

Proof. Suppose, towards a contradiction, that there exists a point q on po, other than p itself which lies on ℓ. Since Γ has no inflections at p the arc po is disjoint from ℓ near p (by Taylor's theorem). Thus we may assume that q is the "first" point of po after p which intersects ℓ, assuming that Γ is oriented so that o is the "end point". After a rigid motion we may assume that ℓ coincides with the x axis, the tangent vector of Γ at p (which is consistent with the orientation of Γ) is parallel to the positive direction of the x-axis, and Γ lies above the x-axis near p, see Figure 7. Now it follows from Lemma 3.1 that q lies to the "left" of p (with respect to the

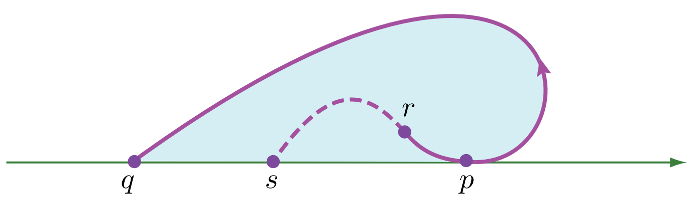


FIGURE 7.

positive directions of the x-axis). Indeed, if q where to lie on the "right" of p, then the inverse of the Beltrami projection applied to the subarc pq of Γ would yield a spherical curve with at least one inflection in its interior according to Lemma 3.1—a contradiction. So connecting the end points of the subarc pq with a straight line yields a simple closed curve C which is $\mathcal{C}^1$-regular in the complement of q. Now note that if a point r of Γ precedes p and is sufficiently close to p, then r lies in the interior of the region bounded by C, because by assumption Γ lies above the x-axis near p. So the arc $o'r$ of Γ, where o' is the initial point of Γ, must intersect C at some point, because o' lies on ∂B, the boundary of the ball associated to the spiral

Γ, which lies outside C. Since Γ is simple, $o'r$ can intersect C only in the interior of the line segment pq. Let s be the last point of Γ prior to r which lies on pq. Then s lies to the left of p, and again it follows from Lemma 3.1 that sp contains an inflection, which is the contradiction we were seeking. $\square$

The last lemma yields:

Lemma 5.10. *Let Γ be a spiral with vertex o. Then either the tangent lines of Γ converge to a line at o, or else Γ intersects every ray emanating from o infinitely often. In particular, if interior of Γ is disjoint from some ray emanating from o, then Γ is $\mathcal{C}^1$-regular.*

Proof. Let L denote the (possibly infinite) length of Γ, $\gamma\colon [0,L) \to \Gamma$ be a unit speed parametrization of $\Gamma - \{o\}$ with $\lim_{t\to L}\gamma(t) = o$, and $T := \gamma'\colon [0,L) \to \mathbf{S}^1$ be the corresponding tantrix. Then we may write $T(t) = (\cos(\theta(t)), \sin(\theta(t))$, for some continuous function $\theta\colon [0,L) \to \mathbf{R}$. Since $\Gamma - \{o\}$ has no inflections, $\theta' \neq 0$. For definiteness, we may suppose that $\theta' > 0$, or θ is strictly increasing. Then either θ is bounded above, or it increases indefinitely. If θ is bounded above, then it converges to a limit, and it follows that T may be continuously extended to $[0,L]$. This would show that Γ is $\mathcal{C}^1$-regular. On the other hand, if θ is not bounded above, then T winds around $\mathbf{S}^1$ indefinitely. By Lemma 5.9, $\gamma(t)/\|\gamma(t)\|$ and $T(t)$ can never be parallel. So γ has to wind around o indefinitely as well. $\square$

Now we can complete the proof of Proposition 5.3. Recall that here we are assuming that the double spiral Γ is not convex. So on at least one of the arms of Γ, say Γ_1, the principal normals point outward with respect to the given proper side S. Take a point p of Γ_1 other than o. Then, for some $r > 0$, there exists a circle C of radius r which lies in S and passes through p. Let $q \in op$ be the point closest to o along Γ_1 such that there passes a circle C of radius r through q which lies in S; see Figure 8.

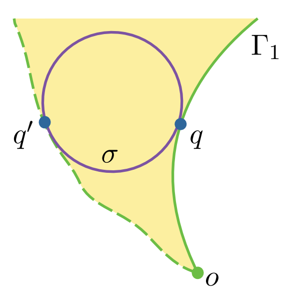


FIGURE 8.

If $q = o$, then each arm of Γ is $\mathcal{C}^1$-regular, by Lemma 5.10. So S has a well-defined angle θ at o. Since by assumption S is a proper side of Γ, $\theta \leq \pi$. On the other hand, if $q = o$, then $\theta \geq \pi$. So we conclude that $\theta = \pi$, which yields that Γ is $\mathcal{C}^1$-regular. Then Lemma 5.8 completes the proof.

So we may suppose that $q \neq o$. Then C must intersect the other arm Γ_2 of Γ at some point (other than o); because, by Lemma 5.9, C does not intersect the arc oq of Γ_1 at any point other than q. Thus if C were disjoint from Γ_2, then we could slide C closer to o along Γ_1 which would be a contradiction.

Let q' be the closest point of Γ_2 to o which intersects C. Then q and q' determine an arc σ in C such that the union of σ with the arc $q'q \subset \Gamma$ forms a simple closed curve which bounds a region outside C. Replacing the arc $q'q$ of Γ with σ then yields a $\mathcal{C}^1$-regular arc $\overline{\Gamma}$. Note that $\overline{\Gamma}$ has at most only two points near which it may not be $\mathcal{C}^2$-regular, namely q and q'. Since $\overline{\Gamma}$ is $\mathcal{C}^1$-regular, however, we may smooth away these singularities by Lemma 5.8, and possibly Lemma 5.5, which completes the proof. Indeed, note that $\overline{\Gamma}$ is semiconvex near q, so removing that singularity will cost precisely one inflection. On the other hand, by Lemma 5.7, $\overline{\Gamma}$ is either convex or semiconvex near q', so removing q' will cost either 0 or 1 inflections respectively. The former case occurs only when Γ is semiconvex, and the latter occurs only when Γ is concave. So the case where Γ is semiconvex will cost 1 inflection, and the concave case will cost 2 inflections, as claimed.

5.2. **Singularities which are double points.** To complete the proof of Theorem 5.1, it remains only to consider the case where p is a singular point of Γ which is also a double point (but is not a multiple singularity). As in Section 5.1, assuming U is sufficiently small, we may identify it with the interior of a ball $B \subset \mathbf{R}^2$ centered at o. Here B will contain precisely two branches or subarcs of Γ, say Γ_1 and Γ_2. Furthermore, Γ_1 is $\mathcal{C}^2$-regular, while Γ_2 has precisely one singularity at o, which is the only point where the two arcs meet. We show that we may perturb Γ_1 and Γ_2 so as to "separate" the singularity and the double point. This perturbation will leave Γ_1 and Γ_2 fixed near their end points, and does not increase the numbers of intersections, singularities, or inflections of these curves. Then we may remove the singularity, which is now simple, via Proposition 5.3, to complete the proof of Theorem 5.1. In summary, all we need is to show:

Proposition 5.11. *Let Γ_i, $i = 1, 2$, be simple arcs which lie in a ball $B \subset \mathbf{R}^2$ centered at o, and intersect each other only at o. Suppose that Γ_1 is $\mathcal{C}^2$-regular, Γ_2 is $\mathcal{C}^2$-regular in the complement of o, and Γ_i have only finitely many inflections. Then there are simple arcs $\widetilde{\Gamma}_i$ in B which coincide with Γ_i near their end points, have the same regularity properties as Γ_i, have no more inflections than Γ_i have, and intersect at most at one point which is not a singularity.*

To establish this proposition, we may assume that Γ_1 and Γ_2 cross each other at o, i.e., one of the arms of Γ_2, say $\overline{\Gamma}_2$, lies on one side of Γ_1 while the other arm lies on the opposite side of Γ_1. Indeed if Γ_2 lies entirely on one side of Γ_1, then we may perturb Γ_1 into the opposite side via Lemma 4.4 which will quickly remove the double point, and leave behind a simple singularity.

Second, since by assumption Γ_i are allowed to have only finitely many inflections, we may assume, after replacing Γ_i by smaller subarcs (and reducing the size of B accordingly), that these curves have no inflections except possibly at o.

Third, we may assume after a rigid motion that Γ_1 is tangent to the x-axis at o, and $\overline{\Gamma}_2$ lies "above" Γ_1 near o; see Figure 9(a). Also note that, by Lemma 5.10, $\overline{\Gamma}_2$ is $\mathcal{C}^1$-regular. Further, after a perturbation we may assume that $\overline{\Gamma}_2$ is not tangent to the x-axis, by the next lemma; see Figure 9(b). More specifically, near o we may represent $\overline{\Gamma}_2$ as the graph of a function f satisfying the hypothesis of the following lemma. Then replacing the graph of f with the graph of the function $\widetilde{f}$ given by this lemma yields the desired perturbation of $\overline{\Gamma}_2$.

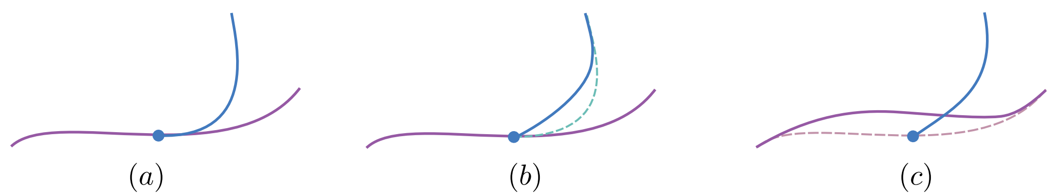


FIGURE 9.

Lemma 5.12. *Let $f\colon [0,a] \to \mathbf{R}$ be a $\mathcal{C}^1$ function which is $\mathcal{C}^2$ on $(0,a]$. Suppose that $f(0) = f'(0) = 0$, and f, $f'' > 0$ on $(0,a]$. Then there exists a $\mathcal{C}^1$ function $\widetilde{f}\colon [0,a] \to \mathbf{R}$, which is $\mathcal{C}^2$ on $(0,a]$, $\widetilde{f}(0) = 0$, $\widetilde{f}'(0) > 0$, $\widetilde{f} \geq f$, $\widetilde{f}'' > 0$ on $(0,a]$, and $\widetilde{f} \equiv f$ near a.*

Proof. Let $\phi\colon [0,a] \to \mathbf{R}$ be a $\mathcal{C}^\infty$ nonnegative function such that $\phi(x) \equiv x$ on $[0,a/3]$, and $\phi \equiv 0$ on $[2a/3,a]$. Set $\widetilde{f} := f + \epsilon\phi$ for constant $\epsilon > 0$. Then $\widetilde{f} \geq f$. Further, $\widetilde{f}'(0) > 0$, $\widetilde{f}'' \equiv f''$ on $(0,a/3] \cup [2a/3,a]$, and $\widetilde{f}'' \to f''$ with respect to the $\mathcal{C}^2$ norm on $[a/3,2a/3]$. So, for small ϵ, we have $\widetilde{f}'' > 0$ on $(0,a]$. □

So we may assume that $\overline{\Gamma}_2$ meets Γ_1 transversely at o. Then by Lemma 4.4 we may perturb Γ_1, by a sufficiently small amount with respect to the $\mathcal{C}^1$-norm, so that it intersects $\overline{\Gamma}_2$ at only one point which is different from o, and does so transversely; see Figure 9(c). This will settle Proposition 5.11, which will in turn complete the proof of Theorem 5.1.

6. Surgery on Regular Double Points

In this section we prove an analogue of the main result of the last section, Theorem 5.1, for multiple points. Recall that, as far as proving Theorem 1.2 is concerned, we may assume that all multiple points of Γ are double points.

Theorem 6.1. *Let M be a Riemannian surface of constant nonnegative curvature, $\Gamma \subset M$ be a $\mathcal{C}^2$-regular closed curve with finitely many inflections and singularities, and p be an isolated double point of Γ which is not a multiple singularity. Then for every open neighborhood U of p which does not contain any other multiple points of Γ, there is a $\mathcal{C}^2$-regular closed curve $\widetilde{\Gamma} \subset M$ such that $\widetilde{\Gamma} = \Gamma$ outside U, $\widetilde{\Gamma}$ has no multiple points inside U, and $\Sigma^+(\widetilde{\Gamma}) \leq \Sigma^+(\Gamma)$.*

By Theorem 5.1 we may assume that p is not a singular point of Γ. Further, as in Section 5.2, we may identify U with the interior of a ball $B \subset \mathbf{R}^2$ centered at o, and let Γ_i, $i = 1, 2$, be the two branches of Γ in B. Also, again as in Section 5.2, we may assume that Γ_1 and Γ_2 cross each other due to Lemma 4.4. Finally note that here Γ_i are $\mathcal{C}^2$-regular, and by the next result we may assume that their point of intersection is not an inflection of either curve:

Proposition 6.2. *Let Γ_i, $i = 1, 2$, be simple $\mathcal{C}^2$-regular arcs which lie in a ball $B \subset \mathbf{R}^2$ centered at o, and intersect each other only at o. Suppose that Γ_i have no inflections except possibly at o. Then there are simple $\mathcal{C}^2$-regular arcs $\widetilde{\Gamma}_i$ in B which coincide with Γ_i near their end points, have no more inflections than Γ_i have, are transversal to each other, and intersect at most at one point which is not an inflection of either curve.*

Assume for now that the above proposition holds. Then after replacing Γ_i with smaller subarcs, and replacing B by a smaller ball centered at the new intersection point of Γ_i, we may assume that Γ_i have no inflections, and intersect ∂B only at their end points. Next note that $B - (\Gamma_1 \cup \Gamma_2)$ consists of precisely 4 components, see Figure 10. Let us call the closure of each of these components a *sector* of $\Gamma_1 \cup \Gamma_2$.

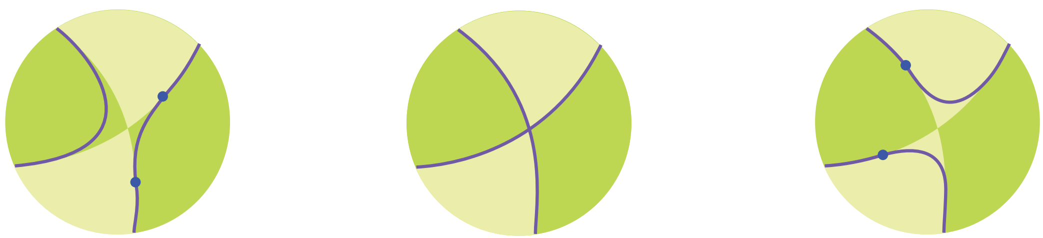


FIGURE 10.

Each sector is bounded in the interior of B by a double spiral which is composed of one arm of Γ_1 and one arm of Γ_2, since by assumption Γ_1 and Γ_2 cross each other. Further, since by assumption Γ_i intersect each other transversely, it follows that each sector is a proper side of the corresponding double spiral, as we defined in Section 5.1. Let us say that a sector is *convex*, *concave*, or *semiconvex* according to whether the corresponding double spiral is convex, concave, or semiconvex respectively, again as defined in Section 5.1. Now pick a pair of opposite sectors, i.e., a pair of sectors which do not share a common boundary spiral. Then there are only two possibilities: either both sectors are semiconvex, or one sector is convex while the other one is concave, as shown in Figure 10. In either case we may use Proposition 5.3 to remove the double point while adding at most two inflections. Finally note that one of these surgeries will leave Γ connected, which completes the proof of Theorem 6.1.

It only remains then to prove Proposition 6.2. First we show that after a perturbation we may assume that Γ_i are transversal. To this end recall that we are assuming that Γ_i cross each other, and suppose that Γ_i are tangent to each other at o. Then after a rotation, we may assume that Γ_i are tangent to the x-axis. Let

$f_i\colon [-a,a] \to \mathbf{R}$ be functions whose graphs coincide with neighborhoods of o in Γ_i. Since Γ_i cross each other, we may assume that $f_1 \geq f_2$ on $[0,a]$, and $f_1 \leq f_2$ on $[-a,0]$. Then the next lemma yields the desired perturbation:

Lemma 6.3. *Let $f\colon [-a,a] \to \mathbf{R}$ be a $\mathcal{C}^2$ function with $f(0) = f'(0) = 0$, and $f'' \neq 0$ on $[-a,0) \cup (0,a]$. Then there exists a $\mathcal{C}^2$ function $\widetilde{f}\colon [-a,a] \to \mathbf{R}$ with $\widetilde{f}(0) = 0$, but $\widetilde{f}'(0) \neq 0$ such that $\widetilde{f}'' \neq 0$ on $[-a,0) \cup (0,a]$, $\widetilde{f} \equiv f$ near $\pm a$, $\widetilde{f} \geq f$ on $[0,a]$, and $\widetilde{f} \leq f$ on $[-a,0]$. Furthermore, if $f''(0) \neq 0$, then $\widetilde{f}''(0) \neq 0$ as well.* □

The proof of the above lemma is very similar to that of Lemma 5.12 and so will be omitted. Now that Γ_i are transversal, we may use Lemma 4.4 to finish the proof of Proposition 6.2. More specifically, we just need to perturb Γ_i to make sure that they do not intersect at inflection points. To this end we may first perturb Γ_1 via Lemma 4.4 to make sure that the intersection of Γ_i is not an inflection of Γ_2. Next we perturb Γ_2, again via Lemma 4.4, near the new intersection point of Γ_i, which will ensure that the intersection of Γ_i is not an inflection of Γ_1 either. This settles Proposition 6.2, and thus completes the proof of Theorem 6.1.

Note 6.4. The basic patterns for the formation of inflections in Figure 10 appear to have been observed by Klein [29, p. 188] in the context of real algebraic curves.

7. Proof of the Main Result

Here we complete the proof of Theorem 1.2. We may assume that the quantities $\mathcal{D}$, $\mathcal{S}$, and $\mathcal{I}$ are all finite. Then Γ is a $\mathcal{C}^2$ curve with a finite number of singularities. Further, the finiteness of $\mathcal{I}$ implies that Γ does not contain any geodesic arcs, and so cannot be contained entirely in a plane passing through the origin. In particular, $\mathrm{conv}(\Gamma)$ has interior points. Recall also that we may assume that all points of Γ which are not simple are double points, for otherwise $\mathcal{D}^+ \geq 3$ and there would be nothing left to prove. Next we show that we may assume that all the antipodal pairs of points of Γ are simple and regular. To this end we need one more fact concerning spherical curves which is a finer version of a lemma of Fenchel [12, 44], since it assumes less regularity.

Lemma 7.1. *Let $\Gamma \subset \mathbf{S}^2$ be a simple arc which is $\mathcal{C}^2$-regular in its interior. If Γ has no inflections in its interior, then it lies in an open hemisphere.*

Proof. Let p be the initial boundary point of Γ, with respect to some fixed orientation, and q be the last point in Γ such that the subarc pq lies in a hemisphere H. Note that pq lies in an open hemisphere, only if q is the final boundary point of Γ, in which case we are done. Suppose then, towards a contradiction, that pq does not lie in any open hemisphere. Then pq must meet the boundary C of H in at least two points. If there are exactly two such points, then they must be antipodal, say $\pm r$. Consider the subarc $\overline{\Gamma} \subset \Gamma$ bounded by $\pm r$. Then $\overline{\Gamma}$ is $\mathcal{C}^1$-regular by Lemma 5.10, hence Lemma 3.1 applies to $\overline{\Gamma}$. Indeed, after rotating C about $\pm r$, we may assume that $\overline{\Gamma}$ is tangent to C at some point. Then Lemma 3.1 implies that $\overline{\Gamma}$ has an inflection in its interior, which is impossible. So pq must meet C in at least 3 points,

say r, s, t, and suppose that s lies in the oriented arc rt. Then rt is again $\mathcal{C}^1$-regular by Lemma 5.10. Furthermore, rt must be tangent to C at s, and it follows once more from Lemma 3.1 that at least one of the subarcs rs or st has an inflection in its interior, which is again impossible. □

Let $\Sigma^+(\Gamma)$ be as in (10), and set

$$\Sigma(\Gamma) := 2(\mathcal{D} + \mathcal{S}) + \mathcal{I}.$$

Lemma 7.1 quickly yields:

Corollary 7.2. *Let $\Gamma \subset \mathbf{S}^2$ be a $\mathcal{C}^2$ closed curve, and $\pm p$ be a pair of antipodal points of Γ. If $\pm p$ are not both regular simple points of Γ, then $\Sigma(\Gamma) \geq 6$ and $\Sigma^+(\Gamma) \geq 4$.*

Proof. Consider the subarcs of Γ with end points $\pm p$. By Lemma 7.1 each of these arcs must have either an inflection, a singularity, or a double point in its interior. So it follows that $\Sigma^+(\Gamma) \geq 2$. Now if either p or $-p$ is a singular or a double point of Γ, then we have $\Sigma^+(\Gamma) \geq 4$ as desired. Further, since $\pm p \in \Gamma$, we also have $\Sigma(\Gamma) \geq \Sigma^+(\Gamma) + 2 \geq 6$, which completes the proof. □

Now we are ready to finish the proof of Theorem 1.2. First note that if Γ is symmetric, then by the last observation we may assume that it has no singularities or double points, because every point of Γ is antipodal to another one. Then Proposition 4.5 completes the proof of the inequality (5). It remains then to prove Inequalities (3) and (4). There are two cases to consider: either o is an interior point of $\mathrm{conv}(\Gamma)$, or o lies on the boundary of $\mathrm{conv}(\Gamma)$.

Suppose first that o lies in the interior of $\mathrm{conv}(\Gamma)$. Then o remains in $\mathrm{conv}(\Gamma)$ after small perturbations of Γ. Thus, using Theorems 5.1 and 6.1, we may remove the singularities and double points of Γ while keeping o in $\mathrm{conv}(\Gamma)$, and without increasing $\Sigma^+(\Gamma)$. Furthermore, $\Sigma(\Gamma)$ will not increase either, since by Corollary 7.2 we may assume that the antipodal points of Γ are all simple and regular. Thus we do not need to perturb Γ near those points (which might risk increasing the number of antipodal pairs of points of Γ). In short, we may assume that $\mathcal{D}^+ = \mathcal{S} = 0$, in which case Proposition 4.5 completes the proof.

It remains now to consider the case where o lies on the boundary of $\mathrm{conv}(\Gamma)$, or equivalently Γ lies in a hemisphere H of $\mathbf{S}^2$. Since $\Gamma \subset H$ and $o \in \mathrm{conv}(\Gamma)$, Γ must intersect the great circle C which bounds H. There are only two cases to consider: either $\Gamma \cap C$ contains a pair of antipodal points, or not.

Suppose first that $\Gamma \cap C$ contains a pair of antipodal points $\pm p$. By Corollary 7.2, $\pm p$ must be simple regular points of Γ. Now using Theorems 5.1 and 6.1, we may remove the singularities and double points of Γ without perturbing $\pm p$, and thus keeping o in $\mathrm{conv}(\Gamma)$. Recall that these operations do not increase $\Sigma^+(\Gamma)$. Furthermore, $\Sigma(\Gamma)$ does not increase either, because all the singularities and double points of Γ are away from $-\Gamma$, by Corollary 7.2, and thus perturbations near these points will not create any new pairs of antipodal points. Now that we have eliminated all the singularities and double points, Proposition 4.5 completes the proof in this case.

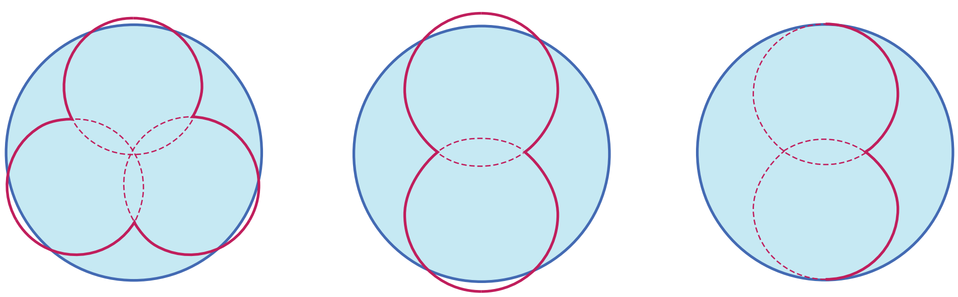
Finally suppose that $\Gamma \cap C$ does not contain any pairs of antipodal points. Then $\Gamma \cap C$ must contain three points p_1, p_2, p_3, which contain o in the relative interior

of their convex hull (this follows from Carathéodory's theorem, as we argued in the proof of Corollary 3.3). Note that at least one of these three points, say p_1, must be a regular simple point of Γ, for otherwise $\Sigma^+(\Gamma) \geq 6$ and we are done. Furthermore, $-p_1 \not\in \Gamma$ by assumption. Now we may use Lemma 4.4 to perturb Γ near p_1, and without increasing $\Sigma(\Gamma)$ or $\Sigma^+(\Gamma)$, so that o lies in the interior of $\mathrm{conv}(\Gamma)$. Indeed, since we are assuming that Γ is not a great circle, there is a point $p_0 \in \Gamma$ which does not lie in the plane of C. So o lies in the tetrahedron T formed by $\{p_0, \dots, p_3\}$. Suppose that C lies in the xy-plane and T lies "above" this plane. Then by construction o lies in the interior of the base of T. So it is clear that if we move p_1 by a small distance below the xy-plane, then o will lie in the interior of T, and therefore in the interior of $\mathrm{conv}(\Gamma)$, as claimed. Since we have already treated this case, the proof is now complete.

8. Sharp Examples

Here we describe some examples which illustrate the sharpness of the inequalities in Theorem 1.2, and consequently establish the sharpness of all the inequalities cited in the introduction.

We begin with inequality (3). Note that there are exactly 10 triples $(\mathcal{D}, \mathcal{S}, \mathcal{I})$ of nonnegative integers such that $2(\mathcal{D}+\mathcal{S})+\mathcal{I} = 6$. Each of these cases may be realized by a spherical curve satisfying the hypothesis of Theorem 1.2. To generate these examples we begin with the curve depicted in Figure 11 (a). This picture illustrates



(a) (b) (c)

Figure 11.

the stereographic projection of a spherical curve from the "North pole" $(0, 0, 1)$ of $\mathbf{S}^2$ into the tangent plane at the "South pole" $(0, 0, -1)$, and the shaded disk corresponds to the "Southern hemisphere". Recall that stereographic projections preserve circles. Thus the curve in Figure 11 (a) which is composed of 3 circular arcs has no inflections ($\mathcal{I} = 0$). Further, it is not hard to see that this curve is disjoint from its antipodal reflection, and so $\mathcal{D} = 0$, since this curve is also simple. Furthermore, this curve contains the origin o in its convex hull, because it intersects the equator in 6 points which form a polygon passing through o. We conclude then that the curve in Figure 11 (a) is an example of type $(0, 3, 0)$. To generate examples of other types, note that we may replace each singularity of this curve with a small loop, see Figure 12. Alternatively, as Figure 12 demonstrates, each singularity may

FIGURE 12.

be smoothed away at the cost of precisely two inflections as we proved in Proposition 5.3. Thus we obtain examples which exhaust all possible 10 cases where equality may hold in (3).

Similarly, one may establish the sharpness of inequality (4). Here there are only 6 cases to consider, since that is precisely the number of possible triplets $(\mathcal{D}^+, \mathcal{S}, \mathcal{I})$ of nonnegative integers such that $2(\mathcal{D}^+ + \mathcal{S}) + \mathcal{I} = 4$. Again each case will be possible. The case $(0, 2, 0)$ is illustrated in Figure 11 (b), and the other cases may be generated by replacing the singularities by small loops, or a pair of inflections as described earlier.

Finally, we consider inequality (5). Here each of the quantities in the triple $(\mathcal{D}^+, \mathcal{S}, \mathcal{I})$ must be even due to symmetry. Thus, for the equality to hold in (5), there are only 3 possible cases: $(0, 0, 6)$, $(0, 2, 2)$, and $(2, 0, 2)$. Again, all these cases may be realized. Figure 11 (c) corresponds to the case $(0, 2, 2)$. This picture depicts one-half of the curve in a hemisphere. Here the depicted arcs are not exactly circular, but have been perturbed slightly so that they have contact of order 2 with the boundary of the hemisphere. Now, gluing this curve to its antipodal reflection yields a $\mathcal{C}^2$-regular closed curve of type $(0, 2, 2)$. The other cases $(2, 0, 2)$ and $(0, 0, 6)$ may again be treated by replacing the singularity in Figure 11 (c) with a small loop or a pair of inflections.

Note 8.1. Each of the spherical curves (or tantrices) we constructed here may be integrated to obtain a closed space curve, since these curves all contain the origin in the interior of their convex hull [17]. Thus we will obtain sharp examples for inequalities (1) and (2). It would be interesting to know if such examples can be constructed in every knot class, or whether the lower bound for these inequalities may be improved according to the topological complexity of the curve. Alternatively, one could also ask whether we may improve or balance these inequalities by adding some geometric terms to their right hand sides. See the paper of Wiener [48] who adapted the Fabricius-Bjerre formula [11, 23] to spherical and space curves. Those results, however, assume that the curves are generic and do not give estimates for the total number of pairs of parallel tangent lines, but only the difference between the number of concordant and discordant pairs.

Note 8.2. If a closed space curve which is generic with respect to the $\mathcal{C}^1$-topology has a pair of discordant parallel tangent lines, then it must have at least two such pairs. This is due to the basic observation that if a simple closed curve $\Gamma \subset \mathbf{S}^2$ crosses its antipodal reflection Γ', i.e., Γ does not lie on one side of Γ', then Γ must have at least two pairs of antipodal points. To see this, suppose that p and p' are

the only antipodal pairs of points of Γ. Let Ω be a region bounded by Γ in $\mathbf{S}^2$ with area at most 2π. Next, let A and B be the two subarcs of Γ determined by p and p'. Then the reflection of one of these arcs, say A, which we denote by A', must lie in Ω, since by assumption Γ' must enter Ω. Now note that the closed curve $A \cup A'$ is simple and symmetric. Thus it bisects $\mathbf{S}^2$. In particular, the region Ω' in Ω which is bounded by $A \cup A'$ must have area 2π. So it follows that $\Omega = \Omega'$, and in particular $\Gamma = \Gamma'$, which contradicts the assumption that Γ crosses Γ'.

Note 8.3. An analogue of Theorem 1.1 for the number of parallel normal lines of a closed space curve may be obtained as follows. If $\gamma\colon \mathbf{S}^1 \to \mathbf{R}^3$ is a $\mathcal{C}^2$ unit speed curve with nonvanishing curvature κ, then its *principal normal* vector field $N\colon \mathbf{S}^1 \to \mathbf{S}^2$, given by $N := T'/\kappa$, is a well-defined spherical curve. Since $T' = \kappa N$, γ has parallel normal lines at a pair of points, if, and only if, T has parallel tangent lines at the corresponding points. Hence Theorem 1.2 may be used to obtain information with regard to the number of parallel normal lines of a space curve. Recall that τ/κ is the geodesic curvature of T. The critical points of this quantity, i.e., the zeros of $(\tau/\kappa)'$, have been called *Darboux vertices* [25]. Now suppose that γ is $\mathcal{C}^4$, let $\mathcal{P}_N$ be the number of (unordered) pairs of points $t \neq s$ in $\mathbf{S}^1$ where normal lines of γ are parallel, i.e., $N(t) = \pm N(s)$, and $\mathcal{V}_d$ be the number of Darboux vertices of γ. We claim that

$$2\mathcal{P}_N + \mathcal{V}_d \geq 6. \tag{11}$$

To establish this inequality, first note that N is the tantrix of T, which is a closed space curve. Thus N contains the origin in its convex hull, and so Theorem 1.2 may be applied to N. Further, assuming that γ is parametrized by arclength, we have $N' = -\kappa T + \tau B$, where $B := T \times N$ is the binormal vector. Thus $\|N'\| = \kappa^2 + \tau^2 \neq 0$, which means that N has no singularities. So, by Theorem 1.2, $2\mathcal{P}_N + \mathcal{I} \geq 6$, where $\mathcal{I}$ is the number of inflections of N. It remains only to note then that the geodesic curvature of N vanishes if and only if $\langle N'', N \times N' \rangle$ vanishes, and a straight forward computation shows that this is the case precisely when $(\tau/\kappa)'$ vanishes. Finally, we should mention that we do not know whether (11) is sharp. A major difficulty in constructing examples here is that, in contrast to the tantrix which has a simple characterization, no complete characterization for the normal spherical image of a closed space curve is known [13].

Acknowledgements

The author thanks Xiang Ma, Gaiane Panina, John Pardon, Bruce Solomon, Serge Tabachnikov, and Masaaki Umehara for useful communications.

References

[1] S. Alexander, M. Ghomi, and J. Wong. Topology of Riemannian submanifolds with prescribed boundary. *Duke Math. J.*, 152(3):533–565, 2010.

[2] S. Angenent. Parabolic equations for curves on surfaces. II. Intersections, blow-up and generalized solutions. *Ann. of Math. (2)*, 133(1):171–215, 1991.

[3] S. Angenent. Inflection points, extatic points and curve shortening. In *Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995)*, volume 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 3–10. Kluwer Acad. Publ., Dordrecht, 1999.

[4] V. Arnold. On the number of flattening points on space curves. In *Sinaĭ's Moscow Seminar on Dynamical Systems*, volume 171 of *Amer. Math. Soc. Transl. Ser. 2*, pages 11–22. Amer. Math. Soc., Providence, RI, 1996.

[5] V. I. Arnold. *Topological invariants of plane curves and caustics*, volume 5 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1994. Dean Jacqueline B. Lewis Memorial Lectures presented at Rutgers University, New Brunswick, New Jersey.

[6] V. I. Arnold. Topological problems in wave propagation theory and topological economy principle in algebraic geometry. In *The Arnoldfest (Toronto, ON, 1997)*, volume 24 of *Fields Inst. Commun.*, pages 39–54. Amer. Math. Soc., Providence, RI, 1999.

[7] K.-S. Chou and X.-P. Zhu. *The curve shortening problem*. Chapman & Hall/CRC, Boca Raton, FL, 2001.

[8] D. DeTurck, H. Gluck, D. Pomerleano, and D. S. Vick. The four vertex theorem and its converse. *Notices Amer. Math. Soc.*, 54(2):192–207, 2007.

[9] M. P. do Carmo and F. W. Warner. Rigidity and convexity of hypersurfaces in spheres. *J. Differential Geometry*, 4:133–144, 1970.

[10] H. G. Eggleston. *Convexity*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 47. Cambridge University Press, New York, 1958.

[11] F. Fabricius-Bjerre. On the double tangents of plane closed curves. *Math. Scand*, 11:113–116, 1962.

[12] W. Fenchel. Über Krümmung und Windung geschlossener Raumkurven. *Math. Ann.*, 101(1):238–252, 1929.

[13] W. Fenchel. On the differential geometry of closed space curves. *Bull. Amer. Math. Soc.*, 57:44–54, 1951.

[14] M. Ghomi. Vertices of closed curves in riemannian surfaces. *To appear in Comment. Math. Helv.*

[15] M. Ghomi. Shadows and convexity of surfaces. *Ann. of Math. (2)*, 155(1):281–293, 2002.

[16] M. Ghomi. Tangent bundle embeddings of manifolds in Euclidean space. *Comment. Math. Helv.*, 81(1):259–270, 2006.

[17] M. Ghomi. h-principles for curves and knots of constant curvature. *Geom. Dedicata*, 127:19–35, 2007.

[18] M. Ghomi. Topology of surfaces with connected shades. *Asian J. Math.*, 11(4):621–634, 2007.

[19] M. Ghomi. A Riemannian four vertex theorem for surfaces with boundary. *Proc. Amer. Math. Soc.*, 139(1):293–303, 2011.

[20] M. Ghomi and B. Solomon. Skew loops and quadric surfaces. *Comment. Math. Helv.*, 77(4):767–782, 2002.

[21] M. Ghomi and S. Tabachnikov. Totally skew embeddings of manifolds. *Math. Z.*, 258(3):499–512, 2008.

[22] M. A. Grayson. Shortening embedded curves. *Ann. of Math. (2)*, 129(1):71–111, 1989.

[23] B. Halpern. Global theorems for closed plane curves. *Bull. Amer. Math. Soc.*, 76:96–100, 1970.

[24] J. Hass and P. Scott. Shortening curves on surfaces. *Topology*, 33(1):25–43, 1994.

[25] E. Heil. Some vertex theorems proved by means of Moebius transformations. *Ann. Mat. Pura Appl. (4)*, 85:301–306, 1970.

[26] S. B. Jackson. Vertices for plane curves. *Bull. Amer. Math. Soc.*, 50:564–478, 1944.

[27] S. B. Jackson. Geodesic vertices on surfaces of constant curvature. *Amer. J. Math.*, 72:161–186, 1950.

[28] F. Klein. Eine neue Relation zwischen den Singularitäten einer algebraischen Curve. *Math. Ann.*, 10(2):199–209, 1876.

[29] F. Klein. *Elementarmathematik vom höheren Standpunkte aus. Dritter Band: Präzisions- und Approximationsmathematik*. Dritte Auflage. Ausgearbeitet von C. H. Müller. Für den Druck

fertig gemacht und mit Zusätzen versehen von Fr. Seyfarth. Die Grundlehren der mathematischen Wissenschaften, Band 16. Springer-Verlag, Berlin, 1968.

[30] A. Kneser. Bemerkungen über die anzahl der extrema des krümmung auf geschlossenen kurven und über verwandte fragen in einer night eucklidischen geometrie. In *Festschrift Heinrich Weber*, pages 170–180. Teubner, 1912.

[31] A. F. Möbius. Über die grundformen der linien der dritten ordnung. In *Gesammelte Werke II*, pages 89–176. Verlag von S. Hirzel, Leipzig, 1886.

[32] S. Mukhopadhyaya. New methods in the geometry of a plane arc. *Bull. Calcutta Math. Soc. I*, pages 31–37, 1909.

[33] R. Osserman. The four-or-more vertex theorem. *Amer. Math. Monthly*, 92(5):332–337, 1985.

[34] V. Ovsienko and S. Tabachnikov. *Projective differential geometry old and new*, volume 165 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2005. From the Schwarzian derivative to the cohomology of diffeomorphism groups.

[35] G. Panina. Singularities of piecewise linear saddle spheres on $\mathbb{S}^3$. *J. Singul.*, 1:69–84, 2010.

[36] H. Rosenberg. Hypersurfaces of constant curvature in space forms. *Bull. Sci. Math.*, 117(2):211–239, 1993.

[37] R. Schneider. *Convex bodies: the Brunn-Minkowski theory*. Cambridge University Press, Cambridge, 1993.

[38] V. D. Sedykh. Four vertices of a convex space curve. *Bull. London Math. Soc.*, 26(2):177–180, 1994.

[39] B. Segre. Global differential properties of closed twisted curves. *Rend. Sem. Mat. Fis. Milano*, 38:256–263, 1968.

[40] B. Segre. Sulle coppie di tangenti fra loro parallele relative ad una curva chiusa sghemba. In *Hommage au Professeur Lucien Godeaux*, pages 141–167. Librairie Universitaire, Louvain, 1968.

[41] B. Solomon. Central cross-sections make surfaces of revolution quadric. *Amer. Math. Monthly*, 116(4):351–355, 2009.

[42] G. Thorbergsson and M. Umehara. A unified approach to the four vertex theorems. II. In *Differential and symplectic topology of knots and curves*, volume 190 of *Amer. Math. Soc. Transl. Ser. 2*, pages 229–252. Amer. Math. Soc., Providence, RI, 1999.

[43] G. Thorbergsson and M. Umehara. Inflection points and double tangents on anti-convex curves in the real projective plane. *Tohoku Math. J. (2)*, 60(2):149–181, 2008.

[44] B. Totaro. Space curves with nonzero torsion. *Internat. J. Math.*, 1(1):109–117, 1990.

[45] R. Uribe-Vargas. On 4-flattening theorems and the curves of Carathéodory, Barner and Segre. *J. Geom.*, 77(1-2):184–192, 2003.

[46] J. L. Weiner. A theorem on closed space curves. *Rend. Mat. (6)*, 8(3):789–804, 1975.

[47] J. L. Weiner. Global properties of spherical curves. *J. Differential Geom.*, 12(3):425–434, 1977.

[48] J. L. Weiner. A spherical Fabricius-Bjerre formula with applications to closed space curves. *Math. Scand.*, 61(2):286–291, 1987.

[49] Y.-Q. Wu. Knots and links without parallel tangents. *Bull. London Math. Soc.*, 34(6):681–690, 2002.

[50] S.-T. Yau. Open problems in geometry. In *Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990)*, pages 1–28. Amer. Math. Soc., Providence, RI, 1993.

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332
E-mail address: ghomi@math.gatech.edu
URL: www.math.gatech.edu/∼ghomi